Genetic Engineering

7

edited by

Peter W. J. Rigby

Genetic Engineering 7

Edited by

Peter W. J. Rigby

Laboratory of Eukaryotic Molecular Genetics,
National Institute for Medical Research,
The Ridgeway, Mill Hill, London

ACADEMIC PRESS · 1988

Harcourt Brace Jovanovich, Publishers
London · San Diego · New York
Boston · Sydney · Tokyo · Toronto

ACADEMIC PRESS LIMITED
24/28 Oval Road
London NW1 7DX

United States Edition published by
ACADEMIC PRESS INC.
San Diego, CA 92101

British Library Cataloguing in Publication Data

Genetic engineering.—Vol. 7
1. Genetic engineering—Periodicals
575.1 QH442

ISBN 0-12-270307-3

Typeset by Paston Press, Loddon, Norfolk
Printed in Great Britain at The Alden Press, Oxford

Preface

The preceding two volumes of this series have considered the major changes which genetic manipulation techniques have brought to cancer biology, immunology and plant science. Another area of eukaryotic molecular biology that has seen changes of equal magnitude is parasitology, a subject of great intrinsic interest and enormous medical importance. The application of modern molecular biological techniques has led to major advances in our understanding of parasitic organisms, to improved diagnostic procedures and to the development of novel routes for vaccination.

In Chapter 1 Christine Clayton discusses the molecular biology of the kinetoplastidae, concentrating particularly on African trypanosomes. She discusses the latest knowledge of some of the peculiar aspects of the biology of these organisms, for example their unusual mitochondria and the phenomenon of antigenic variation, and details our current understanding of the structure and expression of their genes. In the following chapter, John Scaife deals with the application of cloning techniques to the development of novel vaccines, taking the malaria parasite as his major example. These two chapters provide an excellent description of how genetic manipulation procedures can lead to rapid progress in a world-wide infectious disease problem while at the same time providing basic biological information at a level of resolution previously unattainable.

In Chapter 3 Mary Bendig discusses one of the first promised benefits of genetic engineering, the production of large amounts of proteins for human therapeutic use. Initial work in this area concentrated upon the use of *Escherichia coli* expression systems, but it was rapidly realized that the polypeptides produced in such heterologous systems are often incorrectly modified, biologically inactive and difficult to recover. Much effort has therefore gone into the development of eukaryotic expression

systems which obviate these problems. This chapter reviews the systems currently in use and gives a clear overview of the various factors, scientific, commercial and regulatory, which determine the choice of an appropriate system.

As always with books of this type, the Editor must thank the authors for the enormous hard work that they have put into their chapters, and the staff of Academic Press for their effort and encouragement.

Peter W. J. Rigby

London, March 1988

Contents

The cloning of antigen genes from malarial parasites and *Leishmania* species

John G. Scaife

The production of foreign proteins in mammalian cells

Mary M. Bendig

The molecular biology of the Kinetoplastidae

CHRISTINE E. CLAYTON

The Rockefeller University, 1230 York Ave., New York, New York 10021–6399, USA

GENETIC ENGINEERING VOL 7
ISBN 0–12–270307–3

I Introduction

Protozoa of the family Trypanosomatidae are unicellular organisms distinguished by their single flagellum and by a large aggregate of mitochondrial DNA, the kinetoplast. They parasitize plants, vertebrates and invertebrates, causing a wide spectrum of disease, and until recently have been classified mainly on the basis of morphology, host range and disease pathology. The relationship between the organisms that have received most attention from biochemists and molecular biologists is shown in Fig. 1.

Nearly all of the parasites of medical or veterinary importance are transmitted to the mammalian host by biting insects. The salivarian trypanosomes (Mulligan, 1970) multiply in the bloodstream and tissue fluids of mammals. They are included in the blood meals of tsetse flies (Genus, *Glossina*) and replicate in the tsetse alimentary system (the precise position varies with trypanosome species) before infecting a new host via the tsetse saliva during a subsequent feed (Fig. 2). Exceptions to this are *Trypanosoma equiperdum*, which has lost the ability to survive in tsetse and is transmitted venereally, and *Trypanosoma evansi* which is transmitted "mechanically" by flies: in other words, without undergoing any replication in the vector. The salivarian trypanosomes are found in most of Africa, limited only by the distribution of the vector, and so are often known as "African" trypanosomes. Infections can be apparently asymptomatic in game animals, with which the parasite has presumably evolved a stable relationship over many millions of years, but cause dramatic and economically devastating wasting in cattle and an invariably fatal disease (sleeping sickness) in man, with fluctuating fevers, pharmacological disturbance and central nervous system involvement. With these and the other Trypanosomatidae the course of the disease varies widely, depending upon the genetic background of both host and parasite and on other host factors.

Unlike the African trypanosomes, which are exclusively extracellular, the South American trypanosomes (the most important of which is *Trypanosoma cruzi* which causes Chagas' disease) multiply intracellularly in the mammalian host (Fig. 2), damaging the tissue in which they multiply and thereby causing, for instance, heart and arterial disease.

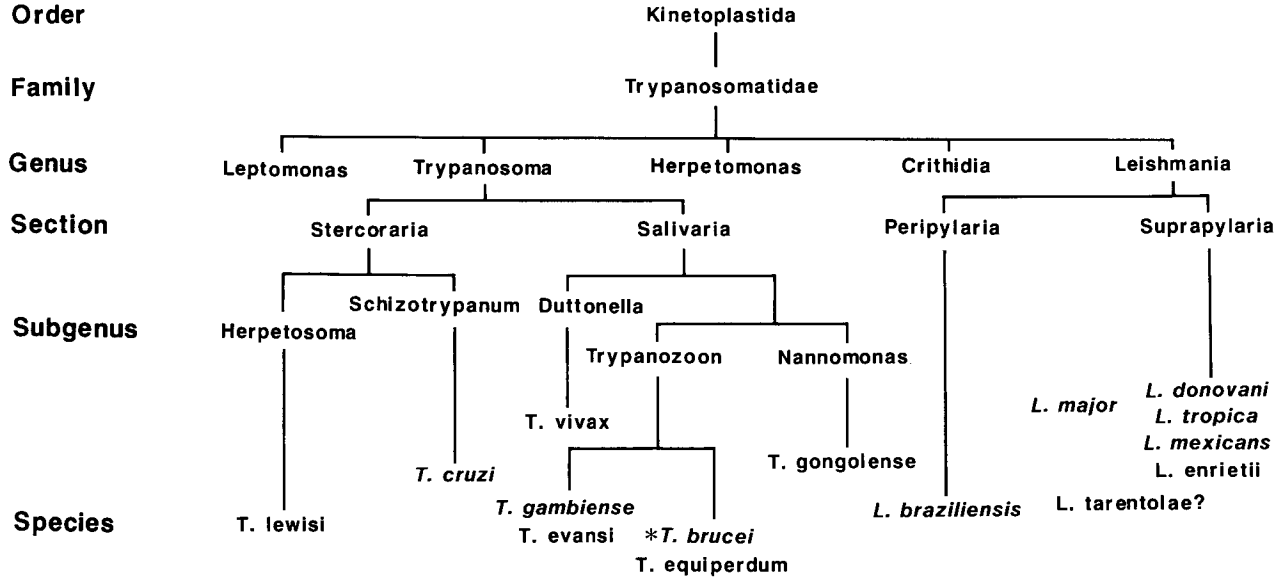

Figure 1 Trypanosomatid classification (information from Hoare (1970) and Chang and Bray (1985)). Human-infective parasites shown in italics. The relationship between *L. major* and *L. tropica*, and the classification of *L. tarentolae*, are somewhat controversial. **T. brucei* is subdivided into two species, *T.b. brucei* and *T.b. rhodesiense*; these are discussed in the text. The diagram takes into account some of the more recent molecular data (e.g. Gibson *et al.*, 1985).

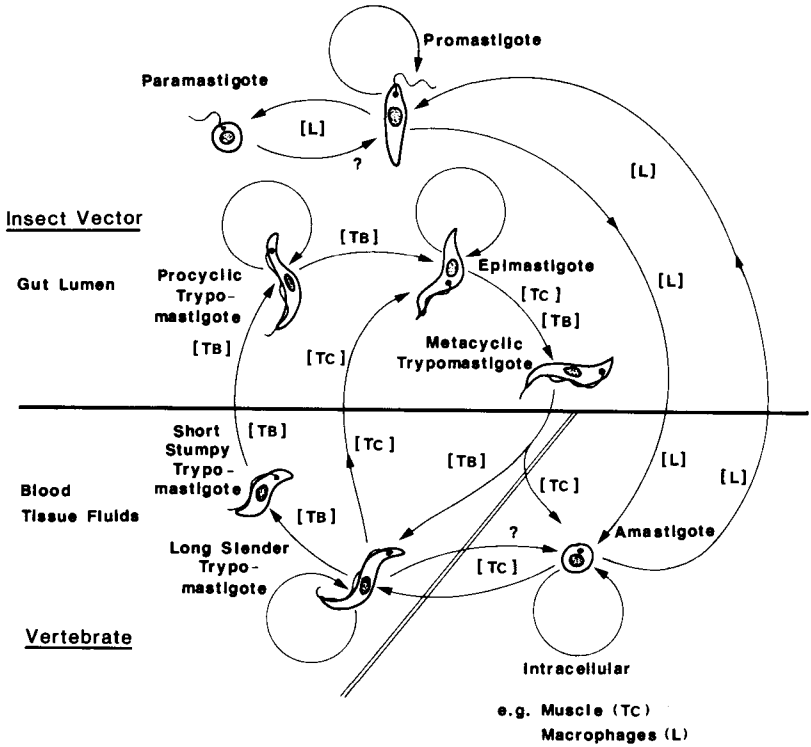

Figure 2 Life cycles of *T. brucei* (TB), *T. cruzi* (TC) and *Leishmania* (L). More details can be found in standard parasitology texts and in papers dealing with developmental regulation (mentioned later in this review).

Infective forms escape into the blood, whence they can infect biting reduvid bugs. Transmission in this case occurs when faeces from the bug contaminate a later bite. The *Leishmanias* (Chang and Bray, 1985) are also intracellular in the mammal, replicating inside macrophages that are taken up from the skin by biting sandflies. Re-infection is via the sandfly proboscis. Leishmaniasis is found in tropical areas throughout the world, with pathology ranging from transient skin sores to extensive destruction of cartilage and fatal inflammation of the viscera, depending on the species of parasite and immunological status of the host.

Drugs for treatment of Leishmaniasis, as for Trypanosomiasis, are generally unsatisfactory, being of moderate to dubious effectiveness and causing severe side effects which necessitate hospitalization—often

impractical in countries with scant resources, sparsely dispersed medical facilities and few roads. To compound the problem, resistance to the available drugs is arising in the parasite, and to insecticides in the vectors.

Each of the parasites mentioned above undergoes a variety of morphological and biochemical changes essential for adaptation to the various environments encountered in the life cycle (Fig. 2). Examples include mitochondrial development and regression, and the growth and loss of the flagellum. The insect forms are all extracellular and relatively easy to culture *in vitro* (some, such as *Crithidia*, will even grow on agar plates), which has facilitated biochemical progress during the past two decades. In contrast, the mammalian forms are more pernickety: the African trypanosomes require either a feeder layer or very carefully adjusted, frequently changed media and relatively low cell densities, while *T. cruzi* and *Leishmania* need host cells in which to grow. Research on the mammalian forms has therefore mainly been restricted to those that will grow to high numbers in laboratory rodents.

The Trypanosomatidae are attractive subjects for the molecular biologist because of the potential applicability of results in treating or preventing a vast amount of tropical disease, and because they exhibit a variety of features of great theoretical interest. The first features to attract the attention of molecular biologists were the unique assemblage of mitochondrial DNA and the phenomenon of antigenic variation in African trypanosomes. Meanwhile, biochemical pathways were investigated with a view to development of drugs. Progress in molecular biological techniques has now expanded the field to include detailed analysis of the regulation of gene expression during the life cycle and attempts to clone genes encoding antigens suitable for expression in prokaryotes and for use as vaccines (see Scaife, this volume, Ch. 2). Recently, the genomes of Trypanosomatidae have been shown to be remarkably unstable and their mode of transcription to be unique. All of these topics combine to make trypanosome research an exciting, varied and fruitful field.

II General features of the genome

A DNA content and ploidy

Estimates of the amount of DNA per cell, from kinetic complexity measurements, have given values of 2.5×10^7 bp per haploid genome for both *Trypanosoma brucei* (Borst *et al.*, 1982) and *T. cruzi* (Lanar *et al.*,

1981) and 5×10^7 bp for *Leishmania* (Leon *et al.*, 1978) with proportions of single-copy DNA varying from 68% in *T. brucei* and 62% in *Leishmania* to 23% in *T. cruzi*. These measurements of genome size are generally confirmed by the frequencies of housekeeping genes in genomic libraries.

The questions of trypanosome ploidy and genetic exchange are only now being resolved after decades of speculation. Mating forms have never been convincingly identified: the only demonstrated mode of replication is by binary fission, so it is impossible to distinguish fusion from division; adhesion organelles specialized for mating have not been seen and the adherence of flagellae *in vitro* (Hughes *et al.*, 1982) is of unknown significance. Feulgen staining of *T. brucei* nuclei suggests a DNA complement of twice that measured by renaturation kinetics (Borst *et al.*, 1982). Characterization and statistical analysis of isoenzyme patterns also indicate that the Trypanosomatidae are diploid and suggest that some form of genetic exchange occurs in populations of *T. brucei*, but not *T. cruzi* or *Leishmania* (Kreutzer *et al.*, 1983; Gibson and Wellde, 1985; Tabayrenc *et al.*, 1986 and reviewed by Tait, 1983). The technique involves native electrophoresis of trypanosome lysates followed by staining for the enzyme activity of interest. Minor differences in amino acid sequence cause slight shifts in electrophoretic mobility. A simple case is illustrated in Fig. 3. Of course, such observations could also be a consequence of post-translational modification or the presence of two genes per haploid genome.

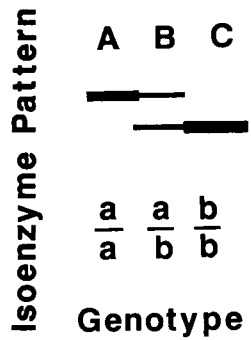

Figure 3 Isoenzyme analysis. Schematic diagram of proteins from strains A, B and C, electrophoresed under native conditions and stained for particular enzyme activity. Strain A is homozygous for a type a; strain C is homozygous for a type b which migrates faster in the gel; strain B is a heterozygote. (For details see Tait (1983).)

The isolation of *T. brucei* housekeeping genes has enabled analysis of restriction enzyme fragment length polymorphisms (RFLPs; see Little, 1981) to be undertaken. This technique is not subject to the reservations noted above. Results confirm diploidy for *T. cruzi* (Gibson and Miles, 1986) and also possible genetic exchange in *T. brucei* (Gibson *et al.*, 1985b), although variant surface glycoprotein genes in the African trypanosomes remain resolutely haploid.

B Genetic experiments with trypanosomes

Attempts to demonstrate recombination between defined genetic markers in the Trypanosomatidae are prejudiced by the lack of any indication as to the stage in the life cycle when this might occur. Glassberg *et al.* (1985) isolated mutants of *Crithidia*, a convenient model because it grows readily *in vitro* in suspension and on agar plates. After mixed growth *in vitro* some recombinants were observed, but unfortunately the frequency was very variable. The phenotypes used in this case were auxotrophies, alterations in colony morphology and resistance to a variety of drugs.

Experiments with African trypanosomes are technically more difficult as they are much more temperamental *in vitro*. However, Jenni *et al.* (1986) recently used analysis of RFLPs to detect hybrid trypanosomes after allowing full development in tsetse flies infected with a mixture of two clones. Mitochondrial DNA in the recombinants was of one parental type only (Sternberg *et al.*, 1986). Reliable establishment of trypanosome mating considerably expands the range of experiments that can be performed and also has considerable epidemiological implications (see Notes added in proof, pp. 55–56).

C Trypanosomatid karyotype

The chromosomes of Trypanosomatidae, like those of yeast, cannot be condensed and spread in the usual way but may be visualized by pulsed-field gel electrophoresis (PFGE), a process involving electrophoresis in an alternating electric field. They then resolve into three broad classes: DNA that remains in the sample slot (owing to constraints either of size or of secondary structure); a group of at least 20 chromosomes ranging in size from 700 to 4000 kb (those around 1000–4000 kb are known as the "megabase" chromosomes); and in *T. brucei, T. equiperdum, T. congolense* and *T. gambiense*, a cluster of minichromosomes of 25–150 kb (Van der Ploeg *et al.*, 1984a). The precise patterns

obtained are dependent on the frequency at which the electric field alternates (see, e.g., Gibson and Borst, 1986). Trypanosomatid karyotypes are in general highly variable, showing extensive differences even between different isolates of a single species (Van der Ploeg *et al.*, 1984a); however, it appears that some chromosomes may be sufficiently stable to allow distinctions to be made reliably between species of *Leishmania* (Giannini *et al.*, 1986). The ribosomal DNA, tubulin, and mini-exon (see Section V), each present in multiple copies, are found in the slot DNA, the megabase DNA, or both, depending on species (Van der Ploeg *et al.*, 1984a); housekeeping genes may also be found on either type (Gibson *et al.*, 1985b). A prominent feature of restriction digests of *T. brucei* DNA is a non-transcribed satellite DNA, consisting of a 177 bp repeat; this has been shown to be located on the minichromosomes (Sloof *et al.*, 1983a,b). This satellite hybridizes with *T. gambiense* DNA, but not with *T. congolense* or with the 196 bp satellite from *T. cruzi*. The genes for the variant surface glycoproteins are not restricted to any one type of chromosome.

Gibson and Borst (1986) studied 14 cloned stocks of salivarian trypanosomes. They found that the size and number of minichromosomes varied markedly between and within species and that the overall karyotype of each stock was unique. Individual chromosomes, identified by cloned DNA probes, also varied in size. Triose phosphate isomerase, calmodulin, aldolase, and one particular variant surface glycoprotein gene "basic" (unexpressed) copy (see section IV) were always in the "slot" DNA, but glyceraldehyde phosphate dehydrogenase and phosphoglycerate kinase genes appeared to be in megabase chromosomes (although a variable amount of hybridization to the slot was seen depending on pulse frequency). These two housekeeping genes were clearly on different chromosomes, the sizes of which were strain-dependent. However, phosphoglycerate kinase was always on the same chromosome as tubulin; occasionally both probes hybridized to two chromosomes, indicating (as there was no increase in copy number) that even homologous chromosomes may differ in length. In *T. cruzi*, all housekeeping genes are found in chromosomes that enter the gel and whose size is isolate-dependent (Gibson and Miles, 1986).

D Molecular biological techniques for classification

Morphologically indistinguishable species of *Leishmania* can produce dramatically different diseases, and the same can be true of the African trypanosomes, so sensitive diagnostic techniques are needed both for treatment and for epidemiological studies. The distributions of isoen-

zyme patterns (reviewed in Tait, 1983; Kreutzer *et al.*, 1983), RFLPs in genomic (Gibson *et al.*, 1985b) and kinetoplast DNA (Borst *et al.*, 1985; Gibson *et al.*, 1985a; Barker *et al.*, 1986; reviewed by Simpson, 1986), and karyotype variations (Giannini *et al.*, 1986; Gibson and Borst, 1986; Paindavoine *et al.*, 1986) have all been correlated to a greater or lesser degree with geographical origin, species and subspecies of isolates. Attempts are also being made to find and use specific repetitive and mitochondrial DNA probes (e.g., Frasch *et al.*, 1983; Barker *et al.*, 1986) that could be used in field situations where sophisticated electrophoretic equipment is not available. The use of specific antibodies is being explored, but this is outside the scope of this review. Among the African trypanosomes, the variant surface glycoprotein gene repertoire is very unstable so VSG probes can be used to distinguish between isolates (Borst *et al.*, 1980; Massamba and Williams, 1984). In most cases, such methods have served to confirm the classic nomenclature but to render it more sophisticated. For instance, it is now generally accepted, from isoenzyme and molecular genetic evidence, that *T. brucei rhodesiense* and *T. brucei brucei* form a continuous spectrum of organisms, rather than separate subspecies as had originally been suspected on the basis of host range. (Only "*T. rhodesiense*" can infect man but both can infect cattle and game; however, recently some individual antigenic variants of *T. brucei* have been found to infect man (Jennings and Urquhart, 1985).) Both can be distinguished at the molecular level from *T. gambiense*, which causes a more chronic disease in man than does *T. rhodesiense* (Gibson *et al.*, 1985a). *Trypanozoon* and *Nannomonas* are clearly more related to each other than to *Duttonella*, with which they were previously thought to be equivalent (Hoare, 1970).

III Structure and function of kinetoplast DNA

A. General

Kinetoplast DNA (kDNA) comprises 10–20% of the total DNA of trypanosomatids. The kinetoplast lies within the single mitochondrion of the organism and is usually found near the base of the flagellum. When spread and viewed in the electron microscope, the kinetoplast DNA looks like a disc that ranges from 5 μm (*T. brucei*) to 2.5 μm (*Crithidia*) in diameter and is made of thousands of catenated circles (the "minicircles") with occasional longer strands (which are actually part of larger "maxicircles") looping out from the edge. The function of minicircles is unknown; the maxicircles contain the classical mitochondrial