MOLECULAR GENETIC ANALYSIS OF PARASITE SURVIVAL IN P. FALCIPARUM MALARIA

ANNUAL REPORT

JEFFREY V. RAVETCH

AUGUST 3, 1992

Supported by

U.S. ARMY MEDICAL RESEARCH AND DEVELOPMENT COMMAND
Fort Detrick, Frederick, Maryland 21702-5012

Grant No. DAMD17-89-Z-9003

Sloan-Kettering Institute
Memorial Sloan-Kettering Cancer Center
1275 York Avenue
New York, NY 10021

Approved for public release; distribution unlimited.

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.
The human malaria parasite continues to pose a significant health problem to one half of the world's population. Entry into endemic areas places those individuals at risk of morbidity and mortality from this protozoan parasite. The lack of effective immunoprophylaxis has refocused efforts at defining the unique biology of this parasite in order to develop more effective chemotherapeutic agents. Work in my laboratory has concentrated on defining the mode of gene regulation in this organism and providing insight into the strategies developed by the parasite to overcome innate host immune responses. The long-term goal of this research is to manipulate the genome of the organism by developing approaches enabling the cloning of parasite DNA and reintroduction of those altered genes back into the parasite's genome. Progress in the past year has been achieved in constructing...
and characterizing yeast artificial chromosome (YAC) libraries from *P. falciparum* and utilizing these sources of stably cloned DNA to define putative promoter elements for asexual intraerythrocytic stage genes. YAC clones have also facilitated the discovery of new transcription units and the relationship of these genes to previously defined chromosomal rearrangements. These studies have resulted in the identification of several DNA sequences which are the sites for nuclear protein binding. These DNA elements have been used in attempts to drive the expression of heterologous reporter genes in DNA transfection studies.
FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

Where copyrighted material is quoted, permission has been obtained to use such material.

Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

Citations of commercial organizations and trade names in this report do not constitute an official Department of the Army endorsement or approval of the products or services of these organizations.

X In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Animal Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

X For the protection of human subjects, the investigator(s) have adhered to policies of applicable Federal Law 45CFR46.

X In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

PI Signature 8/4/92

ConUS/CONF
Nov 94
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Page Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Front Cover</td>
</tr>
<tr>
<td>Report Documentation Page</td>
</tr>
<tr>
<td>Foreword</td>
</tr>
<tr>
<td>Table of Contents</td>
</tr>
<tr>
<td>Body of Report</td>
</tr>
<tr>
<td>Appendix</td>
</tr>
</tbody>
</table>
BODY OF REPORT

BACKGROUND AND SIGNIFICANCE

The complex life-cycle of the malaria parasite results from the ordered expression of specific genes which dictate the pathways of development and differentiation. Despite the central role of these pathways in the physiology of the parasite, little is known of the mechanisms by which gene expression is regulated in this protozoan. Characterization of these pathways requires 1) access to the various stages of the organism, both within its vertebrate and invertebrate hosts, 2) the ability to stably clone and propagate DNA sequences which flank genes of interest and finally, 3) a functional assay for gene regulation. The malaria parasite has proven to pose significant problems for each of these requirements. Although culture methods are available for the asexual, intraerythrocytic stages of *P. falciparum*, gametocyte stages are less reliably obtained in culture and insect stages require infected mosquitoes for isolation of biologically active material. For these reasons, our efforts have focussed initially on the asexual, intra-erythrocytic stages of *P. falciparum*.

Although the successful cloning of a *P. falciparum* erythrocytic stage gene was first described in 1985, it rapidly became apparent that the unusually high (>80%) A+T content of *P. falciparum* DNA was unstable in traditional prokaryotic hosts, resulting in large deletion of sequences, particularly in flanking regions. However, it is precisely these flanking regions where regulatory sequences usually reside for eukaryotic genes that made this technical problem so troubling.

Work over the last two years in the laboratory has focussed on developing approaches for the stable cloning of flanking DNA, the identification of putative promoter regions in that DNA and the development of strategies for the reintroduction and expression of DNA back into the erythrocytic stage parasite. The strategy we have adopted is to utilize the fission yeast, *Saccharomyces cerevisiae*, as a host for the DNA of *P. falciparum* in the form of linear, artificial chromosomes. The rationale behind this approach stemmed from the similarity in A+T content of the DNA of these two organisms and the possibility of using yeast artificial chromosome vectors to construct libraries of *P. falciparum* DNA. As will be described below, this strategy proved to be very successful. With YAC clones for a variety of erythrocytic stage genes in hand, the identification of putative promoter regions has been possible. These clones, which average 100kb in length, have also provided us the opportunity to investigate the distribution of genes along the chromosome by developing a method of transcription mapping. The combination of YAC cloning and transcription mapping has led to the identification of several new genes and the definition of intergenic regions where transcriptional termination and re-initiation has been mapped. These structural studies have provided the DNA targets for functional studies into their role in gene regulation.

One approach to defining the functional role of putative promoter regions has been to determine if these sequences are the sites for specific nuclear protein interactions, indicative of the interaction of transcription factors with their promoters. The interactions can be visualized by electrophoretic mobility shift assays (EMSA) using DNA targets and nuclear protein extracts from infected erythrocytes. Since transcription of many erythrocytic genes is stage-specific, some of the complexes defined by EMSA should reflect this stage specificity.
Protein-DNA complexes have been described in the last year which fulfill these criteria and are likely to represent the interaction of promoters with transcription factors.

Finally, the definition of a promoter is ultimately a functional one, describing a DNA element capable of directing the transcription of a reporter sequence. This approach requires the ability to reintroduce exogenous DNA back into a cell type of interest. In the case of the infected erythrocyte, multiple membrane barriers exist which make this problem more complicated than a typical eukaryotic cell. Successful transformation of an intracellular pathogen has not been achieved as of this writing. The laboratory has been approaching this problem by using a variety of physical and chemical approaches for the re-introduction of DNA into the infected erythrocyte, using transient expression of sensitive reporter genes driven by the putative promoters defined in the preceding paragraph. The results of these experiments will be summarized below.

RESULTS

Construction and characterization of a yeast artificial chromosome library from *P. falciparum*

Molecular genetic studies of the human malaria parasite *Plasmodium falciparum* have been hampered in part due to difficulties stably cloning and propagating parasite genomic DNA in bacteria. This is thought to be a result of the unusual A + T bias (>80%) in the parasite's DNA. Pulsed-field gel electrophoretic separation of *P. falciparum* chromosomes has shown that large chromosomal polymorphisms, resulting from the deletion of DNA from chromosome ends, frequently occur. Understanding the biological implications of this chromosomal polymorphism will require the analysis of large regions of genomic, and in particular telomeric, DNA. To overcome the limitations of cloning parasite DNA in bacteria, we have cloned genomic DNA from the *P. falciparum* strain FCR3 in yeast as artificial chromosomes. A pYAC4 library with an average insert size of approximately 100 kb was established and found to have a three to fourfold redundancy for single-copy genes. Unlike bacterial hosts, yeast stably maintain and propagate large tracts of parasite DNA. Long-range restriction enzyme mapping of YAC clones demonstrates that the cloned DNA is contiguous and identical to the native parasite genomic DNA. Since the telomeric ends of chromosomes are underrepresented in YAC libraries, we have enriched for these sequences by cloning *P. falciparum* telomeric DNA fragments (from 40 to 130 kb) as YACs by complementation in yeast.


Transcription mapping of a 100 kb locus identifying an intergenic region

We have mapped *Plasmodium falciparum* erythrocytic stage transcription units on chromosome 10 in the vicinity of the gene encoding the glycophorin binding protein (GBP130) using yeast artificial chromosomes (YACs). Three erythrocytic stage transcription
units are clustered in a 40 kb region. Two of these genes are closely linked, separated by <2 kb. Nuclear run-on data demonstrate that transcription of these two genes, though unidirectional, is monocistronic. Within this intergenic region are the sites at which transcription of the upstream gene terminates and the GBP13O gene initiates. These studies represent the first description of the minimal and necessary cis-acting elements for transcription termination and initiation in this protozoan parasite.


A putative promoter region is the site of developmentally regulated DNA-protein interactions

The Plasmodium falciparum gene encoding the knob associated histidine-rich protein (KAHRP) is shown to be transcriptionally regulated during its expression in the intraerythrocytic cycle as demonstrated by stage specific nuclear run-on analysis. The genomic organization of the KAHRP gene was determined and the structural basis for the stage specific transcription investigated. A sequence motif with two-fold symmetry was found 160 bp upstream of the RNA initiation site. This sequence element interacts with parasite derived nuclear extracts in a stage specific manner that correlates with the transcriptional activity of the KAHRP gene. These studies suggest a functional role for this structural element in the developmental regulation of a P. falciparum erythrocytic gene.


Development of transfection strategies for P. falciparum

Our efforts during the past year have focussed on introducing exogenous DNA into the infected erythrocyte using electroporation or biolistic particle bombardment. Initial experiments were aimed at determining the efficiency of introduction of exogenous DNA using these approaches, utilizing labelled DNA. DNA fragments were labelled either by radioisotopic or biotinylated nucleotide triphosphate precursors. The tagged DNA could be reproducibly detected in the isolated parasite pellet of the infected erythrocyte 24 hours after its introduction by electroporation of asynchronous parasites. Control uninfected erythrocytes were routinely negative. Greater efficiencies were seen for schizont stage parasites, when compared to ring or trophozoite enriched stages.

A variety of reporter genes were constructed using DNA sequences from the KAHRP gene, the GBP130 intergenic region and the P195 gene upstream sequences 5' of luciferase or chloramphenicol acetyltransferase reporter genes. These constructs were introduced into schizont-stage parasites using the parameters determined from the labelled DNA experiments. Preliminary results from three independent experiments have suggested that the P195 upstream region when coupled to the luciferase reporter gene results in detectable luciferase activity 24 hours after electroporation of schizont stage parasites. No activity was observed for the GBP130 or KAHRP promoters electroporated into schizonts. These results are encouraging
and consistent with the stage specificity of the three promoters utilized since only the P195 is active during the schizont stage. Confirmation of these results using a second reporter, like CAT, is in progress and will be followed by a detailed analysis of the putative promoter element.
CONCLUSIONS AND FUTURE DIRECTIONS

The success of the YAC cloning strategy has facilitated the analysis of gene organization and regulation in the malaria parasite. These approaches will allow for the cloning of complete chromosomes, thereby identifying genetic loci of biological significance, like drug resistance, cytoadherence, invasion and dimorphism using positional cloning strategies. Novel transcription units can be defined and their regulation studied in intact and rearranged chromosomes. The identification of putative promoters and dominant selectable markers using these strategies will be combined to develop methods for stable transformation of *P. falciparum*. Studies completed during the last year have laid the groundwork for these important extensions.


APPENDIX
Characterization of Yeast Artificial Chromosomes From Plasmodium falciparum: Construction of a Stable, Representative Library and Cloning of Telomeric DNA Fragments

DERIK DE BRUIN, MICHAIL LANZER, AND JEFFERY V. RAVETCH

*Department of Molecular Biology, Cornell University Graduate School of Medical Sciences, 1300 York Avenue, New York, New York 10021; and †DeWitt Wallace Research Laboratory, Division of Molecular Biology, Sloan-Kettering Institute, 1275 York Avenue, RS921, New York, New York 10021

Received March 30, 1992; revised July 6, 1992

Molecular genetic studies of the human malaria parasite Plasmodium falciparum have been hampered in part due to difficulties stably cloning and propagating parasite genomic DNA in bacteria. This is thought to be a result of the unusual A - T bias (>80%) in the parasite's DNA. Pulsed-field gel electrophoretic separation of P. falciparum chromosomes has shown that large chromosomal polymorphisms, resulting from the deletion of DNA from chromosome ends, frequently occur. Understanding the biological implications of this chromosomal polymorphism will require the analysis of large regions of genomic DNA, in particular telomeric DNA. To overcome the limitations of cloning parasite DNA in bacteria, we have cloned genomic DNA from the P. falciparum strain FCR3 in yeast as artificial chromosomes. A pYAC4 library with an average insert size of approximately 100 kb was established and found to have three to fourfold redundancy for single-copy genes. Unlike bacterial hosts, yeast stably maintain and propagate large tracts of parasite DNA. Long-range restriction enzyme mapping of YAC clones demonstrates that the cloned DNA is contiguous and identical to the native parasite genomic DNA. Since the telomeric ends of chromosomes are underrepresented in YAC libraries, we have enriched for these sequences by cloning P. falciparum telomeric DNA fragments (from 40 to 130 kb) as YACs by complementation in yeast.

INTRODUCTION

The human malaria parasite Plasmodium falciparum infects more than 200 million people and claims over 2 million lives, annually. P. falciparum is an obligate intracellular protozoan parasite that requires two hosts, a human and a mosquito, for the completion of its life cycle. In the human host the parasite is haploid and multiplies asexually within hepatocytes and erythrocytes through mitotic divisions. The erythrocytic stages can be cultured in vitro (Trager and Jensen, 1976), providing a source of parasites for study. The sexual phase of the life cycle begins when the mosquito host ingests parasite-laden blood from an infected human. During this phase recombination and independent assortment of chromosomes occurs (Walliker et al., 1987). The complexity of the parasite's life cycle and the difficulties in creating and manipulating mutants have severely restricted the use of classical genetic tools for study of this organism.

Pulsed-field gel electrophoresis studies have revealed that the parasite's haploid genome of $3 \times 10^6$ bp contains 14 chromosomes that range in size from 600 kb to 3.5 Mb (Van Der Ploeg et al., 1985; Wellem et al., 1987). Striking polymorphisms have been observed between homologous chromosomes of different geographical isolates. The variations in chromosome size range from 50 to 300 kb, which in some cases corresponds to 15% of the total length of a chromosome (Ravetch, 1989). These chromosomal polymorphisms appear to be the result of large deletions of DNA from chromosome ends. A pathway described in P. falciparum that leads to chromosomal polymorphisms involves a process of chromosome breakage followed by the healing of the breakpoint through the de novo addition of telomere repeats (Pologe and Ravetch, 1986, 1988; Pologe et al., 1990). This chromosome instability and polymorphism reflects the plasticity of the parasite's genome.

The genomic plasticity of P. falciparum is also indicated by the extensive strain-dependent variations seen in antigenic determinants and protein isoforms (see Kemp et al., 1990 for review). Further antigenic differences arise from chromosomal polymorphisms that delete specific antigen genes (Pologe and Ravetch, 1986, 1988; Pologe et al., 1990). We have been interested in determining what additional roles chromosomal polymorphisms play in the parasite's biology and in studying the overall organization and structure of P. falciparum chromosomes. Questions concerning the parasite's genome require the analysis of large fragments of genomic DNA.
DNA in general and of telomeric fragments in particular.

Molecular biological studies on P. falciparum have been hampered by difficulties in stably maintaining large genomic clones of parasite DNA in bacterial hosts (Kochan et al., 1986; Weber, 1988). The problem of clone stability is thought to be a result of the parasite's unusually A + T-rich DNA. The A + T content averages 82% and can approach 90 to 95% in intergenic regions (Pollack et al., 1982; Weber, 1988). As a means of obtaining complete P. falciparum genomic clones, we report here the construction and characterization of a stable, representative yeast artificial chromosome (YAC) library and the YAC cloning of Plasmodium telomeric DNA fragments by complementation in yeast. The parasite offers a unique eukaryotic system in which to study chromosomal stability and structure, and these YAC clones provide necessary reagents for the analysis of the P. falciparum genome.

MATERIALS AND METHODS

Preparation of parasite DNA. Fifty 10-cm plates of the P. falciparum strain FCR3 (ATCC 30932) were grown to a parasitemia of 10% (approximately 2.3 x 10^6 parasites) as described by Trager and Jensen (1976). DNA was prepared after lysis of the infected erythrocytes by digestion with RNase A and Proteinase K, followed by extraction with buffered phenol:chloroform (1:1) in 50-ml plastic tubes and 1481. CAT CTT TAC CTT

10% trichloroacetic acid. The DNA was slowly mixed by inversion. After centrifugation, the DNA was removed by draining it through a needle hole in the bottom of the tube. The aqueous phase was transferred to a fresh tube and extracted again with phenol-chloroform followed by two extractions with chloroform:isoamyl alcohol (24:1). The DNA was then dialyzed extensively against TE (10 mM Tris (pH 8.0), 1 mM EDTA (pH 8.0)).

Digestion of DNA. For YAC library construction, parasite DNA was partially digested with EcoRI at five different concentrations of enzyme (0.15, 0.10, 0.05, 0.025, and 0.01 U/µg DNA) as described in Sambrook et al. (1989). For the Plasmodium telomere cloning, parasite DNA was digested to completion with XhoI YAC vectors pYAC4 (Burke et al., 1991). Digestion of DNA was performed by digestion of the cell suspension with MicroBrew (New England Biolabs, Beverly, MA) and used to inoculate 25 ml of fresh media. The cells were then grown for 7.5 days in 125 ml of media to reach cell densities of 1 x 10^9/ml. This suspension was then used to inoculate 25 ml of fresh media. The cells were then grown for 4 days in 125 ml of media. The YAC library was then screened for clone stability by PCR analysis.

Preparation of parasite DNA. Fifty 10-cm plates of the P. falciparum strain FCR3 (ATCC 30932) were grown to a parasitemia of 10% (approximately 2.3 x 10^6 parasites) as described by Trager and Jensen (1976). DNA was prepared after lysis of the infected erythrocytes by digestion with RNase A and Proteinase K, followed by extraction with buffered phenol:chloroform (1:1) in 50-ml plastic tubes and 1481. CAT CTT TAC CTT

10% trichloroacetic acid. The DNA was slowly mixed by inversion. After centrifugation, the DNA was removed by draining it through a needle hole in the bottom of the tube. The aqueous phase was transferred to a fresh tube and extracted again with phenol-chloroform followed by two extractions with chloroform:isoamyl alcohol (24:1). The DNA was then dialyzed extensively against TE (10 mM Tris (pH 8.0), 1 mM EDTA (pH 8.0)).

Digestion of DNA. For YAC library construction, parasite DNA was partially digested with EcoRI at five different concentrations of enzyme (0.15, 0.10, 0.05, 0.025, and 0.01 U/µg DNA) as described in Sambrook et al. (1989). For the Plasmodium telomere cloning, parasite DNA was digested to completion with XhoI YAC vectors pYAC4 (Burke et al., 1991). Digestion of DNA was performed by digestion of the cell suspension with MicroBrew (New England Biolabs, Beverly, MA) and used to inoculate 25 ml of fresh media. The cells were then grown for 7.5 days in 125 ml of media to reach cell densities of 1 x 10^9/ml. This suspension was then used to inoculate 25 ml of fresh media. The cells were then grown for 4 days in 125 ml of media. The YAC library was then screened for clone stability by PCR analysis.

Preparation of parasite DNA. Fifty 10-cm plates of the P. falciparum strain FCR3 (ATCC 30932) were grown to a parasitemia of 10% (approximately 2.3 x 10^6 parasites) as described by Trager and Jensen (1976). DNA was prepared after lysis of the infected erythrocytes by digestion with RNase A and Proteinase K, followed by extraction with buffered phenol:chloroform (1:1) in 50-ml plastic tubes and 1481. CAT CTT TAC CTT

10% trichloroacetic acid. The DNA was slowly mixed by inversion. After centrifugation, the DNA was removed by draining it through a needle hole in the bottom of the tube. The aqueous phase was transferred to a fresh tube and extracted again with phenol-chloroform followed by two extractions with chloroform:isoamyl alcohol (24:1). The DNA was then dialyzed extensively against TE (10 mM Tris (pH 8.0), 1 mM EDTA (pH 8.0)).

Digestion of DNA. For YAC library construction, parasite DNA was partially digested with EcoRI at five different concentrations of enzyme (0.15, 0.10, 0.05, 0.025, and 0.01 U/µg DNA) as described in Sambrook et al. (1989). For the Plasmodium telomere cloning, parasite DNA was digested to completion with XhoI YAC vectors pYAC4 (Burke et al., 1991). Digestion of DNA was performed by digestion of the cell suspension with MicroBrew (New England Biolabs, Beverly, MA) and used to inoculate 25 ml of fresh media. The cells were then grown for 7.5 days in 125 ml of media to reach cell densities of 1 x 10^9/ml. This suspension was then used to inoculate 25 ml of fresh media. The cells were then grown for 4 days in 125 ml of media. The YAC library was then screened for clone stability by PCR analysis.

Preparation of parasite DNA. Fifty 10-cm plates of the P. falciparum strain FCR3 (ATCC 30932) were grown to a parasitemia of 10% (approximately 2.3 x 10^6 parasites) as described by Trager and Jensen (1976). DNA was prepared after lysis of the infected erythrocytes by digestion with RNase A and Proteinase K, followed by extraction with buffered phenol:chloroform (1:1) in 50-ml plastic tubes and 1481. CAT CTT TAC CTT

10% trichloroacetic acid. The DNA was slowly mixed by inversion. After centrifugation, the DNA was removed by draining it through a needle hole in the bottom of the tube. The aqueous phase was transferred to a fresh tube and extracted again with phenol-chloroform followed by two extractions with chloroform:isoamyl alcohol (24:1). The DNA was then dialyzed extensively against TE (10 mM Tris (pH 8.0), 1 mM EDTA (pH 8.0)).

Digestion of DNA. For YAC library construction, parasite DNA was partially digested with EcoRI at five different concentrations of enzyme (0.15, 0.10, 0.05, 0.025, and 0.01 U/µg DNA) as described in Sambrook et al. (1989). For the Plasmodium telomere cloning, parasite DNA was digested to completion with XhoI YAC vectors pYAC4 (Burke et al., 1991). Digestion of DNA was performed by digestion of the cell suspension with MicroBrew (New England Biolabs, Beverly, MA) and used to inoculate 25 ml of fresh media. The cells were then grown for 7.5 days in 125 ml of media to reach cell densities of 1 x 10^9/ml. This suspension was then used to inoculate 25 ml of fresh media. The cells were then grown for 4 days in 125 ml of media. The YAC library was then screened for clone stability by PCR analysis.
Burke and Olson (1991) with the exceptions that lyticase, 0.1 M sodium citrate, 50 mM EDTA, and 15 mM DTT were replaced with 2.5 ug/ml Zymolyase 20T (ICN Immunobiologicals), 1 M sorbitol, 10 mM sodium phosphate (pH 7.5), 10 mM EDTA, and 30 mM 2-mercaptoethanol.

Mapping YAC clones and pulsed-field gel electrophoresis. Parasite DNA was prepared for PFGE analysis by embedding the infected erythrocytes in an equal volume of 1.25% Incert Agarose (FMC) in TE (10 mM Tris (pH 8.0) and 1 mM EDTA) and then aliquoting 100 ul/block. The embedded parasites were then lysed by adding the blocks to ESP (0.5 M EDTA (pH 8.5), 1% N-lauryl sarcosine, and 2 mg/ml Proteinase K) and incubating at 55°C for 36 h. YACs were grown under selective conditions for 2 days, harvested by centrifugation (3000 rpm for 5 min), and washed once with distilled water followed by 1 M sorbitol. The yeast were resuspended in an equal volume of SPEM (1 M sorbitol, 10 mM sodium phosphate (pH 7.5), 10 mM EDTA, and 30 mM 2-mercaptoethanol) and Zymolyase 20T added to a final concentration of 5 ug/100 ul of cell suspension. The cells were incubated for 1-2 h at 37°C, embedded in an equal volume of 1.25% Incert agarose, and lysed in ESP as described above.

Prior to digestion of the PFGE, the agarose blocks were dialyzed extensively against ET (50 mM EDTA, 1% sorbitol, and 10 mM Tris (pH 8.0) with 1 mM PMSF and 1% Proteinase K) and incubated at 55°C for 36 h. YACs were grown under selective conditions for 2 days, harvested by centrifugation (3000 rpm for 5 min), and washed once with distilled water followed by 1 M sorbitol. The yeast were resuspended in an equal volume of SPEM (1 M sorbitol, 10 mM sodium phosphate (pH 7.5), 10 mM EDTA, and 30 mM 2-mercaptoethanol) and Zymolyase 20T added to a final concentration of 5 ug/100 ul of cell suspension. The cells were incubated for 1-2 h at 37°C, embedded in an equal volume of 1.25% Incert agarose, and lysed in ESP as described above.

RESULTS

YAC Library Construction

The P. falciparum strain FCR3 (Jensen and Trager, 1978) was chosen as the DNA source for the YAC library because the strain is well characterized and is readily available to other researchers. Parasite DNA for the YAC library was prepared in aqueous form and care was taken to minimize shearing (see Materials and Methods). The size of the undigested DNA was approximately 650 kb, and partial digestion with EcoRI yielded fragments that ranged in size from 30 to 400 kb (data not shown). The parasite DNA was cloned into the EcoRI site of the YAC cloning vector pYAC4. Transformants were selected on uracil- and tryprophan-deficient sorbitol agar plates. Random clones were picked and tested for inserts by color assay (Burke et al., 1987). Of 60 colonies assayed, 52 were positive, indicating that approximately 86% of the primary transformants contained inserts.

Random clones were analyzed by PFGE to determine the size range of the inserts. After Southern blotting the PFGE, the membrane was probed with a pBR322 fragment specific for the YAC vector left arm. Figure 1 shows a typical PFGE containing 15 random YACs that range in size from 30 to 290 kb. The largest YAC clone found in the library is 350 kb (data not shown), and the average clone size is approximately 100 kb. To establish a library, primary transformants were picked and color was assayed for inserts. Of 1300 transformants assayed, 1056 positive clones were transferred to 96-well microtiter plates (see Materials and Methods). Since the size of the P. falciparum genome is 3 x 10^7 bp and the average size of the YACs is 100 kb, the library is predicted to have a redundancy of between 3 and 4 for single-copy genes.

Representation of the pYAC4 Library

A PCR-based strategy was used to identify individual genes present in the library (see Materials and Methods). The 11 primary pools of DNA (each pool represents all clones from a single microtiter plate) were screened using oligonucleotide primers specific for the nine loci (representing six different chromosomes).
TABLE 1
Representation of Loci in the YAC Library

<table>
<thead>
<tr>
<th>Locus</th>
<th>Chromosome</th>
<th>No.</th>
<th>Size (kb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KAHRP</td>
<td>2</td>
<td>5</td>
<td>80, 75, 50, 40, 30</td>
</tr>
<tr>
<td>PFMDR1</td>
<td>5</td>
<td>5</td>
<td>N/D</td>
</tr>
<tr>
<td>CSP</td>
<td>3</td>
<td>4</td>
<td>110, 100</td>
</tr>
<tr>
<td>CARP</td>
<td>N/D</td>
<td>4</td>
<td>280, 190, 180, 145</td>
</tr>
<tr>
<td>SERA</td>
<td>2</td>
<td>3</td>
<td>170, 125, 55</td>
</tr>
<tr>
<td>P-195</td>
<td>9</td>
<td>3</td>
<td>75, 60, 55</td>
</tr>
<tr>
<td>MSA2</td>
<td>2</td>
<td>2</td>
<td>170, 150</td>
</tr>
<tr>
<td>GBP-130</td>
<td>10</td>
<td>2</td>
<td>110, 60</td>
</tr>
<tr>
<td>EBA-175</td>
<td>4</td>
<td>1</td>
<td>50</td>
</tr>
</tbody>
</table>

Note: KAHRP, knob-associated histidine-rich protein; PFMDR1, P. falciparum multidrug resistance gene 1; CSP, circumsporozoite protein; CARP, clustered asparagine-rich protein; SERA, serine repeat antigen; P-195, merozoite surface antigen 1 (MSA1) or precursor of major merozoite surface antigen (PMMSA); MSA2, merozoite surface antigen 2; GBP-130, glycoserin binding protein-130; EBA-175, erythrocyte binding antigen-175. Chromosome, parasite chromosome: No., number of primary DNA pools positive for the locus by PCR; Size, size of YAC clone isolated by a second round of PCR using secondary pools of DNA as determined by PFG. Only two of the four primary YAC clones were isolated for the CSP locus. The sizes of the PFMDR1 YACs were not determined (N/D). The CARP gene has been reported to be located on Chromosome 2 (Wellems et al., 1991). However, probes to CARP and probes recovered from the ends of the CARP YACs hybridize not to Chromosome 2, but to a much larger undetermined chromosome (N/D).

Integrity and Stability of Plasmodium YAC Clones

The YAC clones from a representative region of parasite DNA, the P-195 locus (known as tmsa1 or PMMSA; Holder et al., 1985; Mackay et al., 1985), were chosen for detailed mapping studies to determine whether the YAC-cloned DNA was identical to the native parasite DNA. Figure 2 shows a typical PFG mapping experiment for the three P-195 YACs identified. The YACs are 75, 60, and 55 kb in size, and when restriction enzyme maps of the YACs are compared to FCR3 genomic DNA (Fig. 2), the fragments are identical, indicating that no obvious rearrangements have occurred. Additional long-range restriction mapping of YACs from two other loci, the GBP-130 (Lanzier et al., 1992) and the KAHRP (de Bruin et al., manuscript in preparation), gave similar results.

Since the major concern with P. falciparum genomic clones is stability, the behavior of six YAC clones in long-term cultures was examined (Fig. 3). Two YACs, each from the GBP-130 locus (Fig. 3A), the KAHRP locus (Fig. 3B), and the P-195 locus (Fig. 3C), were monitored for gross rearrangements in size over 25, 50, and 75 consecutive generations of growth in selective media. When Southern blot analysis using gene-specific probes against PFG-separated YACs was performed, no obvious changes were detected over this time course. In addition, a comparative HindIII restriction analysis between the two P-195 YACs and genomic FCR3 parasite DNA revealed no smaller rearrangements of the YAC-cloned DNA during propagation (Fig. 3D). Together the data presented in Figs. 2 and 3 show that unlike parasite genomic clones in bacterial hosts, P. falciparum DNA is stable in yeast.

YAC Cloning P. falciparum Telomeric DNA Fragments

Due to the unique structure of the ends of chromosomes, these sequences are underrepresented in standard YAC libraries. Our interest in the telomere-proximal regions of P. falciparum chromosomes required a method for obtaining large telomeric DNA fragments. The successes in cloning human telomeres by comple-
mentation in yeast using YAC-based systems (Riethman et al., 1989; Cross et al., 1989; Brown, 1989) suggested that similar methods for cloning *P. falciparum* telomeric regions could be used. Since the parasite telomere repeat (GGGTT(T/C)A) is similar to the human repeat sequence (GGGTTA), a YAC-based approach to isolating telomeric regions seemed feasible.

After *NheI* restriction enzyme digestion of FCR3 DNA, the sizes of the DNA fragments that hybridize with an oligonucleotide probe specific for the *P. falciparum* telomere repeat sequence (OL-PFTel) are between 30 and 200 kb (data not shown). *NheI*-digested parasite DNA was ligated with a 50X molar excess of the YAC vector pJS97. After transformation into yeast, mitotic stability of the linearized vector requires complementation by parasite sequences that can function as substrates for a putative yeast telomerase.

To enrich for large linear clones and remove any plasmid background, the behavior of the transformants when plated onto 5-fluoro-orotic acid agar plates was observed. FOA is a pyrimidine analog that is toxic to yeast if they contain the *ura3* gene, which encodes an enzyme required for uracil biosynthesis (Boeke et al., 1984). *ura3* cells are resistant to FOA and the loss of *URA3* from a yeast can be monitored as growth on FOA. When *URA3* is contained on a plasmid, the frequency of loss is much higher than that of a wildtype chromosomal

![Figure 3](image-url)

**FIG. 3.** Stability of *P. falciparum* YAC clones in culture. YAC clones from loci on three different chromosomes, the YFP/G152 and YFPP/G12. A), the KAHRP (YPF/G13 and YFPH/G12. B), and the P-195 (YPF/G15 and YFPH/G14, C), were grown from the frozen master plates in selective media. PFG sample blocks were prepared from the yeast at the initial time point (Lane 1) and at time points representing 25, 50, and 75 consecutive generations from the initial inoculation. (A, B, C) PFG-separated YACs (15 to 20 s over 18 h at 180 V) that have been Southern blotted and hybridized with gene-specific probes. (D) A comparison between *HindIII* digests of the P-195 YACs and FCR3 DNA (PF). The digests were separated as above, and Southern analysis was performed using the P-195-specific probe.

![Figure 4](image-url)

**FIG. 4.** Analysis of *P. falciparum* telomere YAC clones. Nine putative parasite telomere YACs were PFG separated ramped pulse time from 5 to 25 s over 22 h at 180 V and analyzed by Southern analysis: (A) YACs hybridized to a probe specific for the parasite telomere repeat (OL-PFTel) and (B) YACs hybridized to a probe for the *Plasmodium* subtelomeric repeat sequence, PFrep20. The *BglII* exonuclease sensitivity of one parasite telomere YAC clone (TS) is shown in C. Yeast genomic mini-prep DNA from TS was treated with *BglII*. Aliquots were removed at the indicated time points and the reaction was stopped by the addition of EGTA. Samples were digested with *SacI* prior to Southern analysis using the OL-PFTel probe.

**URA3**. This frequency is reflected in the degree of papillation of the yeast when plated onto FOA. Our experiments showed that if the *URA3* is contained on a large (>35 kb) *Plasmodium* YAC clone, the degree of papillation is either equivalent to or slightly (two to three times) higher than that seen for the yeast chromosomal *URA3*. In addition, short (<20 kb) linear plasmids (and presumably the smaller YAC clones) behave as circular plasmids and papillate at much higher frequencies (data not shown). These differences may result from the greater mitotic stability of longer linear clones (Murray et al., 1986).

Putative *Plasmodium* telomere YACs were picked from the primary transformants in top agar and replica plated onto uracil-deficient and FOA agar plates. The degree of papillation of the transformants relative to the papillation of a *ura3* yeast strain, a *URA3* strain, and a strain harboring a plasmid copy of *URA3* (pJS97) was assayed. Transformants that either failed to grow or papillated at low frequency were picked from the uracil minus replica and transferred to microtitre plates for further analysis. The clones were transferred to a nylon filter for colony screening and hybridized against the parasite telomere OL-PFTel probe. Of the 192 clones examined, 66 (34%) hybridized to this probe.

Figure 4A shows Southern analysis of a PFG containing nine putative parasite telomere clones hybridized with the OL-PFTel probe. These clones range in size from 40 to 130 kb. Although equivalent amounts of DNA were loaded on the gel, differences in the intensity of hybridization to OL-PFTel are seen. This may result from variations in the number of parasite telomere repeats present in the clones. To determine if the parasite telomere sequences were at the end of the clones, the
hybridization to OL-PFTel was examined after treatment with BamHI exonuclease. Figure 4C shows an example clone, T8, after digestions with BamHI and RalI. Southern blotting and hybridization with OL-PFTel. The diffuse, smeared band seen at time zero is characteristic of DNA fragments from chromosome ends that are associated with variable lengths of telomere repeats. The rapid disappearance of the parasite telomere hybridizing sequence relative to a more telomere-distal sequence, the URA3 gene on the pJS97 vector (data not shown) demonstrates that this repeat sequence is at the end of the YAC clone.

The chromosome end origin of these Plasmodium telomere repeat sequences was determined by examining the clones for hybridization to parasite sequences that are found exclusively in subtelomeric regions. The PFprep20 sequence is a 21-bp repeat that has been found to be associated only with the telomeric DNA fragments of P. falciparum chromosomes (Patarapotikul and Langsley, 1988). In addition, PFprep20 sequences can be lost from either end of a chromosome as a result of the deletion events that lead to chromosomal polymorphisms (Kemp et al., 1990). Figure 4B shows that five of the nine Plasmodium telomere YAC clones are positive by hybridization to PFprep20. The data presented in Fig. 4 suggest that true parasite telomere fragments have been cloned. However, these data cannot rule out the possibility that some internal telomere repeat-like sequences may also have been cloned in this experiment.

The DNA sequences adjacent to the Nhel cloning site of the T8 telomere clone were recovered in E. coli and used as a probe against Southern-blotted, PFG-separated, parasite chromosomes. A fragment of approximately 800 bp was end rescued, and this probe hybridizes to a single large parasite chromosome (possibly Chromosome 10) (Fig. 5). Parasite strains from three different geographical regions, Gambia (FCR3), Honduras (HB2), and New Guinea (D10), are shown in Fig. 5, and hybridization with this end probe demonstrates a typical chromosomal polymorphism.

**DISCUSSION**

Despite being a major health threat to most of the world's population, little is known about the molecular biology of P. falciparum. The unusual plasticity of the parasite's genome offers a unique system in which to study chromosome structure, stability, and polymorphism. Essential methods of molecular biology, such as the cloning of large, stable, tracts of parasite DNA in bacteria, the ability to transfect exogenous DNA into P. falciparum, and the in vitro culture of the entire parasite life cycle, currently are unavailable to researchers. Furthermore, classical genetic analysis of the parasite is unfeasible on a routine basis due to the complex life cycle of P. falciparum, which alternates between vertebrate and invertebrate hosts. Toward overcoming these limitations by providing a stable source of cloned genomic parasite DNA, we have constructed a representative genomic library from P. falciparum. Initial studies indicated that parasite DNA could be stably propagated in a yeast host (Triglia and Kemp, 1991). This report details the construction of a stable, representative, and unarranged yeast artificial chromosome library from P. falciparum and the YAC cloning of parasite telomeric DNA fragments.

A pYAC4 library of 1056 clones with an average insert size of 100 kb has been established. Since only 300 clones of 100 kb each are needed to represent the entire 30-Mb parasite genome once, the Plasmodium pYAC4 library constructed was predicted to have a three- to fourfold redundancy for single-copy genes. Of the nine loci examined in this report, all were represented in this library by at least one clone and the majority (six out of nine) were present either at or above the expected number. In addition, screening the pYAC4 library with probes that are not associated with any known coding region has shown that unlike bacterial libraries, even intergenic sequences are represented in the YAC library at expected numbers (data not shown). Long-range restriction enzyme mapping of P. falciparum YACs has found no obvious differences between the restriction enzyme maps of the YAC-cloned DNA and that of the native genomic parasite DNA (this paper; Lanzer et al., 1992). Data on the long-term stability of six independent YAC clones monitored over 75 generations also support the observation that yeast can stably maintain large tracts of parasite DNA.

The ends of P. falciparum chromosomes appear unstable and are frequently deleted both in the wild and during in vitro culture. The loss of telomeric regions results in chromosomal polymorphisms through a mechanism of chromosome breakage and healing by de novo telomere addition at the breakpoint. Polymorphisms have been studied in detail for parasite Chromosomes 1, 2, and 8 (Pologe and Ravetch, 1986, 1988; Pologe et al., 1990). However, it is not yet known whether such telomeric deletions can occur at the ends of all parasite
chromosomes or if some chromosomes are resistant to these events. Furthermore, the biological implications of such events remain unclear. To study the structure of the chromosome ends and the mechanism responsible for generating polymorphisms, we have cloned parasite telomeric DNA fragments by complementation in yeast. Of nine Plasmodium telomere clones examined in detail, all were positive by hybridization to parasite telomere repeat sequences. This hybridization was sensitive to Bal31 exonuclease activity indicating that the parasite telomere repeats were located at the end of the YAC clone. In addition, five of these nine were positive by hybridization to the P. falciparum repeat sequence PFrep20, which is located exclusively in subtelomeric regions. Not all parasite chromosomes contain PFrep20 because breakage and healing events that generate polymorphisms delete these sequences.

These YAC clones will not only provide important reagents for studying the organization and structure of the P. falciparum genome, but will also help in the mapping and positional cloning of new loci. We have probed YAC clones from the GPB-130 locus using PCR-amplified, labeled cDNAs as probes (transcription unit mapping) and have identified two new erythrocytic stage genes (Lanzer et al., 1992). Mapping data established a tight linkage (<3 kb) between the GBP-130 and a newly identified locus called the 3.8 gene. Nuclear run-on analysis of the intergenic region has shown that it contains minimal regulatory elements required for transcription initiation and termination. This is a vital first step in the development of a parasite transfection system.

Together the pYAC4 library clones and the Plasmodium telomere YACs aid in constructing contig maps for genetically defined loci, such as the chloroquine resistance locus on Chromosome 7 (Wellems et al., 1991), and for entire parasite chromosomes. Chromosomal polymorphism through breakage and telomere healing frequently occurs at the Kahrp locus of Chromosome 2 (Polohe and Ravetch, 1986, 1988). Over 300 kb from the end of Chromosome 2 is deleted from some parasite strains, and we have recovered this telomeric region of the chromosome from FCR3 as contiguous YAC clones. With the YAC clones as reagents and using the transcription unit mapping methods to derive probes, we can begin to examine the polymorphic regions of parasite chromosomes for novel sequences and new genes. The data obtained will help in understanding what is unique about the genome organization of the parasite and why it has evolved such diverse complexity.

ACKNOWLEDGMENTS

This work was supported by U.S. Army Grant DAMD 17-89-Z-9033 and the World Health Organization. J.V.R. is supported by the Burroughs-Wellcome Fund, and M.L. is supported by fellowships from the Deutscher Akademischer Austauschdienst and from the World Health Organization. We thank Phil Hieter for the YAC vectors, the yeast strains, and his advice on YAC cloning. Gordon Langley for his gift of the PFrep20 probe. Kim Lee Sim for the EBA-175 PCR primers. Diane Wirth for the PFMDR1 PCR primers, and Arthur Lustgut and Mary Ann Oslev for their helpful discussions on yeast genetics and methods. Expert technical assistance was provided by J. Carter Ralph, Mark Samuels, and Jahan Dadgar. The library is available for screening upon request by contacting one of the authors (J.V.R.) at the Memorial Sloan-Kettering Cancer Center.

REFERENCES


Kochan, J., Perkins, M., and Ravetch, J. V. (1986). A tandemly re-
peared sequence determines the binding domain for an erythrocyte receptor binding protein of *P. falciparum*. Cell 44: 699-696.

Lanzer, M., de Bruin, D., and Ravetch, J. V. (1992). Transcription mapping of a 100 kb locus of *Plasmodium falciparum* identifies an intergenic region in which transcription terminates and initiates.

EMBO J. (in press).


### Genom 1755 - (70) - C

<table>
<thead>
<tr>
<th>Generation</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>FF12</td>
<td>G012</td>
</tr>
<tr>
<td>100</td>
<td>1234</td>
<td>5678</td>
</tr>
<tr>
<td>94</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Genom 1755 - 4 - (65) - Com

<table>
<thead>
<tr>
<th>Generation</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>150</td>
<td>IF6</td>
<td>IF4</td>
</tr>
<tr>
<td>100</td>
<td>1234</td>
<td>5678</td>
</tr>
<tr>
<td>94</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Legend:**
- FF12
- G012
- IF6
- IF4
- P-195
- 75
- P-195 - Hind III
Genom 1755 - 5 - (G)

FCR3
HB2
D10

Mb
1.9 -

1.64 -
Transcription mapping of a 100 kb locus of *Plasmodium falciparum* identifies an intergenic region in which transcription terminates and reinitiates

Michael Lanzer, Derik de Bruin and Jeffrey V. Ravetch

Dwight Wallac Research Laboratory, Division of Molecular Biology, Sloan Kettering Institute, 1275 York Avenue, New York, NY 10021, USA

Communicated by P Borst

We have mapped *Plasmodium falciparum* erythrocytic stage transcription units on chromosome 10 in the vicinity of the gene encoding the glycophorin binding protein (GBP130) using yeast artificial chromosomes (YACs). Three erythrocytic stage transcription units are clustered in a 40 kb region. Two of these genes are closely linked, separated by <2 kb. Nuclear run-on data demonstrate that transcription of these two genes, though unidirectional, is monocistronic. Within this intergenic region are the sites at which transcription of the upstream gene terminates and the GBP130 gene initiates. These studies represent the first description of the minimal and necessary cis-acting elements for transcription termination and initiation in this protozoan parasite.

Key words: malaria parasite promoter RNA processing SV40 enhancer yeast artificial chromosomes

Introduction

The protozoan parasite responsible for the most severe form of human malaria, *Plasmodium falciparum*, alternates between vertebrate and invertebrate hosts. During this complex life cycle gene expression is regulated, as indicated by the accumulation of stage-specific transcripts (Ravetch et al., 1985; Pologe and Ravetch, 1986; Waters et al., 1989; Wesseling et al., 1989). The mechanisms regulating gene expression in this important human pathogen are largely unknown, due in part to the difficulties of cloning and stably maintaining potential regulatory sequences in standard prokaryotic vectors and hosts. Frequent deletion and rearrangement of *P. falciparum* DNA has been observed in *Escherichia coli* hosts (Kochan et al., 1986; Wellens and Howard, 1986; Weber, 1988). This may result from the extreme A + T content of the parasite's genome, which is -80% overall and approaches 90% in non-coding regions (Gorman et al., 1982; Pollack et al., 1982). Thus, our knowledge of the *P. falciparum* genome has been largely restricted to short and isolated fragments of the coding region, with little information on the organization of genes or the elements that regulate transcription. Defining these elements would help in our understanding of the mechanisms regulating gene expression and host switching. Furthermore, a basic understanding of the structural elements involved in transcriptional processes is a necessary first step for the development of a transfection protocol for *Plasmodium*.

Large fragments of *P. falciparum* DNA have been cloned and propagated as artificial chromosomes in yeast (Triglia and Kemp, 1991; de Bruin, D., Lanzer, M. and Ravetch, J. V., manuscript in preparation), suggesting that DNA from this parasite can be stably maintained in the yeast host. YAC clones spanning a 100 kb region of the GBP130 locus were isolated and erythrocytic stage transcripts were mapped. Two additional transcription units were identified flanking the GBP130 gene. Using nuclear run-on assays, these transcripts were shown to be monocistronic. Sequence analysis revealed that the transcripts are continuous with their DNA. By mapping the termination and initiation sites for these genes a short intergenic region has been identified in which the minimal sequence elements required for these processes must be contained. A structural motif within this intergenic region reveals homologies to another plasmodial upstream region, suggesting common elements involved in transcriptional processes of these genes.

Results

Clustering of blood stage genes on chromosome 10 in the vicinity of the GBP130 gene

A *P. falciparum* YAC library was constructed by cloning genomic DNA, partially digested with EcoRI, into the YAC vector pYAC4 (de Bruin, D., Lanzer, M. and Ravetch, J. V., manuscript in preparation). YAC clones containing the GBP130 gene were identified by PCR analysis using oligonucleotides derived from the GBP130 coding region. Two YAC clones, designated FF12 and GC12, with insert sizes of 100 and 50 kb, respectively, were obtained. The two YAC clones were mapped with several restriction enzymes, including BamHI, NcoI and HindIII. The restriction analysis reveals that the YAC clone GC12 is contained within clone FF12. When compared with total *P. falciparum* genomic DNA, the YAC clones were found to be unarranged (Figure 1A). These clones have been stably propagated over 50 generations.

To determine the location of additional erythrocytic stage genes surrounding the GBP130 gene, a transcription map was derived. DNA was prepared from yeast cells harboring the GBP130 YACs, digested with the appropriate restriction enzymes, size-fractionated by pulse-field gel electrophoresis and transferred to nitrocellulose. As a control, DNA from untransformed yeast cells was prepared and treated accordingly. The nitrocellulose filter was probed with radiolabelled total cDNA which was prepared from poly(A)+ RNA isolated from erythrocytic stage parasites. To increase the hybridization signals the cDNA was amplified by PCR using GC rich, random primers. This choice of primers favored the amplification of coding sequences in *P. falciparum*. Since the distribution of GC rich sequences varies, cDNA species are amplified unequally. Therefore, the intensity of hybridization signals does not necessarily correlate with RNA accumulation or RNA stability (see Figures 3 and 4 for comparison). Hybridization
Fig. 1. Structural organization of the GBP130 YAC clones. (A) Restriction mapping of *P. falciparum* genomic DNA (lanes marked P1). GBP130 YAC clone DNA (lanes marked FF12 and GC12, respectively) and yeast DNA (lane marked S.c.) were digested with the restriction endonucleases indicated, size-fractionated by pulse-field gel electrophoresis and transferred to nitrocellulose. The nitrocellulose filter was probed with a cDNA clone to the GBP130 gene. A DNA size standard is indicated. (B) Transcription mapping. GBP130 YAC clone DNA (FF12 and GC12) and yeast DNA (S.c.) were digested with *N*ol and *Hind*III, respectively, size-fractionated by pulse-field gel electrophoresis and transferred to nitrocellulose. The nitrocellulose filter was hybridized with a radiolabeled total cDNA probe. The genomic DNA was generated from RNA isolated from an asynchronous erythrocytic culture of *P. falciparum*. Hybridization signals specific for plasmidial sequences are identified as GBP130, 3.8 and X1. Additional hybridization signals evident in the autoradiograms were disregarded since they also appear in the yeast control lane. (C) Genomic organization and restriction map of the GBP130 locus. The two GBP130 YAC clones, FF12 and GC12, are indicated. Shaded rectangles indicate the location of transcription units. The precise location of the X1 transcription unit was not determined as denoted by the jagged borders (*N*, *N*ol; B, *Bam*HI; H, *Hind*III).

Signs for the *N*ol and *Hind*III digests are shown in Figure 1B. The patterns were related to the restriction map thereby defining a chromosomal transcription map. Bands that were also present in the control lane marked S.c. were disregarded. In addition to the GBP130 gene at least two new erythrocytic transcription units were identified and designated as 3.8 and X1. The chromosomal location of these transcription units is shown in Figure 1C.

Fig. 2. Genomic organization and the sequence of the 3.8 gene and the GBP130 intergenic region. (A) Genomic organization and clones. Open reading frames are indicated by rectangles. Several genomic and cDNA clones are shown. A triangle in the genomic clone 2044 indicates an internal deletion generated during cloning in *E. coli* (Kochan et al., 1986). (B) Sequence of the GBP130 intergenic region. The intergenic region is flanked by the 3.8 and GBP130 open reading frames as indicated. Two polyadenylation sites for the 3.8 gene are underlined. A dur location of 305 bp is indicated by large boxes. A sequence element with homology to the SV40 core enhancer sequence is highlighted. The GBP130 transcription start site is indicated by an arrowhead.

**Two blood stage transcription units are tightly linked**

Restriction mapping of this locus revealed that two of these blood stage genes, the 3.8 and the GBP130 genes, are tightly linked by a short intergenic region of < 2 kb. To define this intergenic region the locus was cloned and sequenced, as presented in Figure 2. The sequence reveals the presence of two open reading frames, separated by a 3 kb region of AT rich sequence, characteristic of non-coding sequence in *P. falciparum*. Probes were derived from the 5' open reading frame and used to isolate cDNA clones from a library.
generated from asynchronously growing erythrocytic stage parasites (Ravetch et al., 1985). A comparison of the cDNA sequence with the genomic sequence revealed an intron of 201 bp, which is flanked by consenus acceptor and donor sites. Thus, the linkage of two blood stage genes as deduced from the transcription map is confirmed by these structural data.

The erythrocytic stage-specific expression of these two genes was determined by Northern analysis. Total cellular RNA was isolated from the ring, trophozoite and schizont intra-erythrocytic forms of the parasite. When the Northern blot was hybridized with a probe to the 3.8 gene, a single RNA species of 3.8 kb was observed in ring and trophozoite stage parasites (Figure 3). Rehybridization of the same blot with a GBP130 probe revealed the GBP130 transcript of 6.6 kb in trophozoites (Figure 3). A probe from the intergenic region did not hybridize to any RNA species (data not shown).

Transcription of the GBP130 gene is monocistronic and continuous

A nuclear run-on analysis was performed to determine whether transcription of the 3.8 and the GBP130 genes are monocistronic or polycistronic. If transcription of these two genes is monocistronic then the intergenic region should contain regulatory signals. Nuclei were isolated during the trophozoite stage, in which both genes are transcribed. Preformed transcription complexes were allowed to elongate in the presence of labelled nucleotides. The radiolabelled nascent RNA was used as a probe for DNA fragments spanning this locus (Figure 4). Fragment size and base composition were approximately equivalent for these fragments. Nascent RNA hybridized to fragment 1 which contains the 3.8 gene and to fragments 3 - 6 which span the GBP130 gene. By contrast, the intergenic region fragment 2, did not hybridize to nascent RNA, indicating that it is not transcribed (Figure 4). Thus, the 3.8 and the GBP130 genes are transcribed independently in a monocistronic fashion. The 3.8 kb transcript is terminated with an efficiency of >90% as calculated from the ratio of radioactivity bound to fragment 1 versus 2.

The precise termination site for the 3.8 transcript was determined by RNase protection experiments (Figure 5A). A single stranded, radiolabelled RNA probe complementary to the 3.8 mRNA was generated (probe A, Figure 5C) and hybridized to poly(A)+ trophozoite RNA. Upon RNase digestion a major species of 300 bp was detected, as well as two minor species of 130 and 140 bp in size. The major species maps to the consensus polyadenylation site (AATAA) at position 1500 (see Figure 2B), while the minor species map to the polyadenylation site at position 1270. The polyadenylation site at position 1500 has been confirmed by the isolation of poly(A) containing cDNA clone (A228) which has utilized this site. These data verify the orientation of the 3.8 transcript and its termination site in the intergenic region.

The 5' end of the GBP130 gene was determined by S1 mapping and primer extension (Figure 5B). The primer extended product was recovered from the gel and analyzed by anchored PCR, confirming that the primer used hybridized to and extended the GBP130 RNA. Both primer extension and S1 analysis map the 5' end of the GBP130 RNA to position 3216 (numbering refers to Figure 2B).
M. Lanzer, D. de Bruin and J. V. Ravetch

Fig. 5. Mapping of termination initiation sites in the intergenic region. (A) Termination of the 3.8 gene. RNase protection analysis of single stranded, radio-labelled probe specific for the 3.8 gene (probe A, position 1170 - 1582 in Figure 2B) was generated and hybridized to 10 µg of poly(A)-trophophagic RNA. Upon RNase digestion products were analyzed by gel electrophoresis (lane marked P1). The sites of the products were compared with a standard. A control using yeast poly(A)-RNA was analyzed in parallel (lane marked yeast). (B) Initiation of the GBP130 gene. SI mapping analysis: a single stranded, end labelled probe (probe B, position 9209 - 9321 in Figure 2B) specific for the GBP130 gene, was hybridized to 10 µg of total cellular RNA prepared from trophophocytes. Upon digestion with SI ribonuclease and 3 mg/ml of enzyme for 90 min products were analyzed by gel electrophoresis. The size of the product indicated by an arrow was compared with a sequencing reaction. Primer extension analysis: an end labelled primer corresponding to position 3263 - 3280 in Figure 2B was hybridized to 1 µg of poly(A)-trophophagic RNA. Extension products were analyzed by gel electrophoresis and compared with a sequencing reaction of genomic DNA using the same primer. The primer extended product was recovered from the gel, amplified by anchored PCR technology (Godzik et al., 1986), cloned and sequenced. (C) Schematic drawing of the locus. The probes used for RNase protection assay (probe A) and for SI mapping probe B are indicated. The termination site for the 3.8 transcript is indicated by a hexagon and the initiation site for the GBP130 gene by the arrowhead.

Fig. 6. a-amanitin sensitive transcription of the GBP130 gene. Nuclei were isolated from trophophytic parasites. One aliquot of the nuclei preparation was incubated with 100 µg/ml of a-amanitin prior to transcription. A gene expressed only during the resistant stage, the CS gene (Rusin et al., 1983) and the ribosomal RNA genes (Langley et al., 1983) were analyzed in parallel for comparison.

Transcription of the GBP130 gene is sensitive to the RNA polymerase inhibitor a-amanitin as determined by nuclear run-on analysis (Figure 6).

Fig. 7. Sequence analysis. Sequence elements derived from the GBP130 intergenic region and from the upstream region of the P.knowlesi CS gene are compared with the SV40 core enhancer region.

poly d/C [µg] 0 1 2 3 4 5 6
poly d/C

Fig. 8. Interaction of the GBP130 sequence element with nuclear extracts. 2 µl of double stranded, end labelled oligonucleotides containing the GBP130 sequence element were incubated with 5 µg of crude nuclear extracts derived from asynchronously growing parasitic cultures. The sequence of the oligonucleotide is shown at the bottom. The amount of poly d(C)/d(G) added to the binding assay is indicated. For cross competition experiments 50 ng of unlabelled GBP130 oligonucleotides or 2 µg of DNA fragment containing the GBP130 intergenic region were added. In addition 2 µg of pUC18 DNA and fragments containing the upstream region either of the KAhRP gene (M. Lanzer, D. de Bruin and J. V. Ravetch, manuscript in preparation) or the P105 gene (Mysler, 1990) were tested for their ability to compete. A control experiment was performed using extracts from uninfected erythrocytes (lane marked RBC extracts).

These data indicate that the GBP130 gene is closely linked to another blood stage gene, which is transcribed in the same orientation. Since transcription of the 3.8 and GBP130 genes is monocentric, the region between both genes must contain the minimal elements that signal both the termination and initiation of transcription in P. falciparum blood stage genes.

Structural analysis of the intergenic region

The intergenic region defined above was examined for sequence elements indicative of eukaryotic promoters. The sequence at position 3029 - 3063 (highlighted in Figure 2B) shows homology to the core region of the SV40 enhancer sequence (Weis et al., 1983) and to a sequence element found in the upstream region of the Plasmodium knowlesi CS gene (Rusin et al., 1983) (Figure 7). To determine whether this sequence element interacts with nuclear proteins, gel retardation assays were performed (Figure 8). Oligonucleotides containing this element were incubated with nuclear extracts derived from asynchronously growing P. falciparum erythrocytic cultures. A stable complex is observed, even in the presence of high concentrations of non-specific competitor DNA. The stability of this complex was analyzed by cross competition experiments. Neither pUC18 DNA nor DNA fragments containing the upstream region of the KAhRP (M. Lanzer, 1952).
The 3' end of the 3.8 gene was found to map to consensus polyadenylation sites which are flanked by long poly(A) and poly(T) tracks. These sequences have the potential to form stem-loop structures in the transcribed RNA which may be associated with the termination of transcription. This region has features characteristic of termination sites defined for the simian malaria parasite *P. knowlesi* (Ruiz et al., 1987). cDNA clones isolated for the 3.8 gene predict an open reading frame encoding a novel plasmodial protein. Comparison of this sequence with the protein database (Dayhoff, December 1991) revealed homology to the family of serine kinases, particularly in the region between amino acids 80 and 170, the enzymatic active site.

A unique initiation site was observed for the GBP130 gene as well as for two other plasmodial genes (the P195 and the KAHRP genes, M. Lanzer, D. de Bruin and J.V. Ravetch, manuscript in preparation). In contrast, multiple initiation sites have been suggested for the three other plasmodial genes investigated to date (the CS gene of the simian malaria parasite *P. knowlesi* (Ruiz et al., 1987); the P7250 gene of the rodent parasite *Plasmodium yoelii* (Lewis, 1990) and the P195 gene of *P. falciparum* (Myler, 1990). One reason for this difference may be due to the frequent pausing of reverse transcriptase in AT rich regions, which could be misinterpreted as multiple initiation sites. Comparison of genomic and cDNA sequences indicates that the GBP130 gene is continuously transcribed. However, post-transcriptional processing of the transcript occurs through cis-splicing and polyadenylation. Transcription of the GBP130 gene is sensitive to the RNA polymerase inhibitor α-amanitin. Similar to other eukaryotic genes transcribed by α-amanitin sensitive polymerases, the sequences immediately upstream of the initiation site for the GBP130 gene contain features suggestive of eukaryotic promoters. A sequence element in the GBP130 intergenic region was found to be homologous to the core region of the SV40 enhancer (Weber et al., 1983) and to a similar sequence motif in the upstream region of the *P. knowlesi* CS gene (Ruiz et al., 1987). The GBP130 sequence element was found to bind to nuclear proteins derived from erythrocytic stage parasites in a sequence-specific manner in mobility shift assays. Although these homologies are suggestive of promoter elements, the lack of a functional assay for putative plasmodial promoters, either in vitro or in vivo, limits the conclusions that can be drawn regarding the role of this sequence in parasite gene transcription. We would expect that this element is involved in more general transcriptional processes and not in stage-specific regulation, since it is present in genes transcribed at different stages of the parasite's life cycle. Precise stage-specific regulation of the GBP130 gene may be mediated by the large direct duplication that is unique for the upstream sequence of this gene.

Materials and methods

**Cultivation of parasites**

The *P. falciparum* strains A2 and FCR3 were grown and maintained as described by Traeger and Jansen (1976a) and by Traeger et al. (1981). If not stated otherwise, the cloned *P. falciparum* strain A2 was used. Parasitic cultures.
were synchronized by Percoll-density gradient centrifugation (Kurten et al., 1985). No gametocytes were observed in the culture under the growth conditions employed.

**Construction of Pfalciparum YAC library**

A Pfalciparum YAC library was constructed as described (de Bruin, D., Lanzer, M., and Ravetch, J.V., manuscript in preparation). Genomic DNA was prepared from the Pfalciparum strain FCR3 (Gorman et al., 1982), partially digested with EcoRI, and inserted into the EcoRI cloning site of the YAC vector PYAC4 (Burke et al., 1985). Yeast spheroplasts (strain AB1380, ATCC 20843) were transformed with the ligated mixture as described by McCormick et al. (1989) with the exception that polyamines were excluded. Transformants were selected on media lacking either uracil and tryptophan. The YAC library was screened by PCR analysis (Heard et al., 1989; Green and Olson, 1990).

**Mapping of YAC clones**

YAC clones used in agarose plugs (Schwartz, J. and Cantor, 1984) were digested with restriction endonucleases and size fractionated by pulse field gel electrophoresis using a Bio-Rad CHEF-DR II system [pulse field conditions: ramped pulse from 2.5 to 10 s over 18 h at 170 V, 1% agarose gel (FMC), 0.5X TBE, at 14°C]. DNA was transferred to nitrocellulose filters and hybridized with nick translated DNA fragments or with radiolabeled total eDNA. Probes for transcription mapping were prepared from an asynchronous erythrocytic culture [50 mM Tris-HCl pH 8.3, 75 mM KCl, 3 mM MgCl2, 10 mM DTT, 1 mM dATP, 1 mM dCTP, 1 mM dGTP, 1 mM dTTP, 50 ng/μl nick labelled total eDNA, 40 U of renatured DNA (Promega) and 600 U of M-MLV H reverse transcriptase (superscript, BRL)] for 60 min at 43°C. The total RNA was purified by column chromatography and amplified by PCR in the presence of [α-32P]dCTP using the TAG-IT kit (BIOS) which uses deoA as primers. Hybridization conditions (Kochan et al., 1989) included 200 ng/ml total yeast RNA as competitor.

**Nuclear run-on analysis**

All steps were carried out on ice. At a parasitemia of ~5% Pfalciparum cultures were chilled on ice. The contents of 30-100 mm Petri dishes were collected and washed once in 1X PBS buffer. Erythrocytes were lysed by the addition of an equal volume of 0.1% saponin (Wallach, 1982), followed by one wash in solution A (20 mM PIPES pH 7.5, 15 mM NaCl, 60 mM KCl, 14 mM c-Mercuriethanol, 0.5 mM EDTA, 0.15 mM spermine, 10 mM spermidine, 0.125 mM PMSF). The parasite pellet was resuspended in 3 ml of solution A and transferred to a dounce homogenizer.

**Preparation of nuclear extracts**

Parasites were prepared by saponin lysis (Wallach, 1982). The following method was adapted from Schreiber et al. (1985). About 5 x 10^7 parasites were resuspended in 1 ml of lysis buffer (10 mM HEpes pH 7.9, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 1 mM DTT, 0.5 mM PMSF and 0.65% NP-40). Nuclei were collected by centrifugation and extracted with 100 μl of extraction buffer (20 mM HEpes pH 7.9, 0.1 M NaCl, 1 mM EDTA, 1 mM EGTA, 1 mM DTT and 1 mM PMSF). After 15 min of vigorous shaking the extract is cleared by centrifugation, yielding a protein concentration of 1-2 μg/ml. 5 μg of crude nuclear extract were incubated with 2 fmol of double-stranded, end labelled oligonucleotides for 20 min at room temperature (20 mM HEpes pH 7.9, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, 5% glyceral, 0.25 mg/ml BSA, 2 μg poly d(C) or as indicated; final volume: 15 μl). Binding assays were analyzed by gel electrophoresis (4% polyacrylamide, 5% glycerol and 0.5× TBE).

**Bacterial strains and libraries**

To minimize recombination and deletion events plasmid DNA was propagated in the E. coli host, SURE (Stratagene). Two libraries of P. falciparum strain A21, a pUC19 plasmid cDNA (Kochan et al., 1986) and a λgt11 genomic library were screened using standard methods (Maniatis et al., 1989). The integrity of all clones and sequences was confirmed by Southern analysis.

**Primer extension**

1.0 pmol (5 x 10^6 c.p.m.) of end labelled oligonucleotide primer [5'-GAACGATCCTCAAGTTAGTATATCGTTGATC-3'] and 1 μg of plasmid DNA were precipitated and hybridized in gel electrophoresis (Kochan et al., 1989). After ethanol precipitation the primer was extended at 43°C for 90 min (50 mM Tris-HCl pH 8.3, 75 mM KCl, 3 mM MgCl2, 10 mM DTT, 1 mM dGTP, 40 U of RNase H (Promega) and 100 U of M-MLV H reverse transcriptase (superscript, BRL)]. Products were analyzed by gel electrophoresis. Primer extension products were recovered from the polyacrylamide gel (Maniatis et al., 1989) and eluted with NaCl and amplified (Loh et al., 1989). Amplified DNA fragments were cloned into pUC18 and sequenced by the universal forward primer.
Transcription units in *P. falciparum*


Received on December 23, 1991, revised on February 18, 1992.
A sequence element associated with the *Plasmodium falciparum* KAHRP gene is the site of developmentally regulated protein-DNA interactions

Michael Lanzer, Derik de Bruin and Jeffrey V. Ravetch*
DeWitt Wallace Research Laboratory, Division of Molecular Biology, Sloan-Kettering Institute, 1275 York Avenue, New York, NY 10021, USA

Received March 16, 1992. Revised and Accepted May 6, 1992

ABSTRACT

The *Plasmodium falciparum* gene encoding the knob associated histidine-rich protein (KAHRP) is shown to be transcriptionally regulated during its expression in the intraerythrocytic cycle as demonstrated by stage specific nuclear run-on analysis. The genomic organization of the KAHRP gene was determined and the structural basis for the stage specific transcription investigated. A sequence motif with two-fold symmetry was found 160 bp upstream of the RNA initiation site. This sequence element interacts with parasite derived nuclear extracts in a stage specific manner that correlates with the transcriptional activity of the KAHRP gene. These studies suggest a functional role for this structural element in the developmental regulation of a *P. falciparum* erythrocytic gene.

INTRODUCTION

Nearly half of the world’s population lives in malaria endemic areas. Transmitted to humans by the bite of an infected mosquito, malaria parasites multiply asexually in hepatocytes then in erythrocytes. The most severe form of human malaria is caused by the protozoan parasite *Plasmodium falciparum*, claiming over three million lives annually. The high mortality associated with *P. falciparum* results from the occlusion of capillaries by infected erythrocytes which adhere to endothelial cells (1, 2, 3, 4). The cytoadherence of infected erythrocytes is dependent upon the interaction of parasite encoded proteins that are translocated to the erythrocytic membrane with receptors expressed on endothelial cells, such as CD36 and ICAM-1 (5, 6). Parasite mutants have been described that exhibit reduced cytoadherence (3, 7, 8). This phenotype was linked to the deletion of a parasite encoded gene, the knob associated histidine-rich protein (KAHRP) (9).

During the asexual intraerythrocytic cycle of parasite development, the KAHRP gene is expressed in a stage specific manner, as indicated by the analysis of steady state RNA (9) and protein accumulation (10, 11). The molecular mechanisms regulating the expression of this important parasite gene are unknown. To dissect the structural motifs associated with the developmental regulation of *P. falciparum* genes, we have studied the expression and regulation of the KAHRP gene during the intraerythrocytic cycle. A comparison of transcriptional activity as determined by nuclear run-on analysis and RNA accumulation revealed that the KAHRP gene is transcriptionally regulated. A sequence element with two-fold symmetry has been found to interact in a stage specific manner with nuclear extracts. This stage specific interaction correlates with the transcriptional activity of this gene, suggesting that this sequence element may be involved in the developmental expression of the KAHRP gene.

MATERIALS AND METHODS

Cultivation of Parasites

The clonal *P. falciparum* strains FCR3-A2 and FVO− were grown and maintained as described (29). Different intraerythrocytic stages were separated by percoll/sorbitol gradient centrifugation (30). This method allows the separation of these stages with great accuracy as determined in blood smears. This is true for the ring and the schizont stage preparation.

Northern analysis

After saponin lysis of infected erythrocytes (31), total cellular RNA was isolated by the acidic guanidinium-phenol chloroform method (32). 1.5 µg of total cellular RNA was fractionated on a 0.8% agarose-formaldehyde gel, transferred to nitrocellulose and hybridized with the nick-translated probe LP20 to the KAHRP gene (9).

Nuclear run-on analysis

All steps were carried out on ice. At a parasitemia of about 5%, cultures were chilled on ice. The contents of thirty 10 cm petri dishes were collected and washed once in 1× PBS buffer. Erythrocytes were lysed by the addition of an equal volume of...
0.1% saponin (31), followed by one wash in solution A (20 mM PIPES pH 7.5, 15 mM NaCl, 60 mM KCl, 14 mM β-mercaptoethanol, 0.5 mM EGTA, 4 mM EDTA, 0.15 mM spermine, 0.5 mM spermidine, 0.125 mM PMSF). The parasite pellet was resuspended in 3 ml of solution A and transferred to a dounce homogenizer. 200 μl of a 10% NP-40 solution was added and six strokes with a B pestle were applied. Nuclei were collected (4,000 rpm for 10 min, Sorvall SM24 rotor) and washed once in solution A. 5 x 10^9 nuclei were transcribed at 37°C for 10 min in 600 μl of solution B (50 mM HEPES pH 7.9, 50 mM NaCl, 10 mM MgCl₂, 1.2 mM DTT, 10 mM creatine phosphate, 1 mM GTP, 1 mM CTP, 4 mM ATP, 25% glycerol, 25 units/ml rRnasin (Promega), 0.2 mg/ml creatine kinase, 0.5 μM [α-35P]UTP 3000C/mmol). Radiolabelled RNA was isolated (32) and purified by TCA precipitation. Typically 2 x 10^7 cpm were incorporated into nascent RNA, with a specific activity of 9 x 10^6 cpm/μg. 1 μg of nascent RNA was hybridized to single stranded DNA fragments (0.2 pmol) immobilized on nitrocellulose. The prehybridization and hybridization conditions were as described (33). Filters were washed three times for 20 min in 2 x SSC, 0.1% SDS at room temperature, twice at 55°C in 0.1 x SSC, 0.1% SDS, followed by one wash at 42°C in 2 x SSC, 0.5 μg/ml of RNase A for one hour. Filters were dried and exposed overnight at 70°C with an intensify screen.

Bacterial strains and libraries

To minimize recombination and deletion events plasmid DNA was propagated in the E. coli host, SURE (Stratagene). A lambda gt11 genomic library of P. falciparum (strain A2) was screened using standard methods (34). The integrity of all clones and sequences were confirmed by Southern analysis.

Primer extension

0.1 pmol (1.5 x 10^5 cpm) of end labelled oligonucleotide primer (5'CATTAATAAATACGATAATATTGAAATTTAC, position 1819 to 1850 in Fig. 3B) and 0.4 μg of poly A+ ring RNA from the parasite strain FT3-A2 were coprecipitated and hybridized (34). After ethanol precipitation the primer was extended at 43°C for 90 min (50 mM Tris/HCl pH 8.3, 75 mM KCl, 3 mM MgCl₂, 10 mM DTT, 1 mM dNTP, 50 units/ml rRnasin (Promega), and 300 units of M-MLV H reverse transcriptase (superScript, BRL)). Products were analyzed by gel electrophoresis. Primer extension products were recovered from the polyacrylamide gel (34), tailed with dGTP and amplified (35). Amplified DNA fragments were cloned into pUC18 and sequenced by using the universal forward primer.

RNase protection assay

A Nsi I EcoRI fragment corresponding to position 1419 to 2215 in Fig. 3B was cloned into pGEM3. A single stranded RNA probe complementary to the KAHRP mRNA was generated, gel purified and hybridized to 1 μg of poly A+ ring RNA from the parasite strain FC3-A2 and FVO respectively. Hybridization and digestion conditions (0.5 units/ml of RNase A, 100 units/ml of RNase T1 for 30 min at 37°C) were followed as recommended by the manufacturer of the ribonuclease protection assay kit (Ambion).

Preparation of nuclear extracts

Parasites were prepared by saponin lysis (31). Nuclear extracts were prepared as described (36). About 5 x 10^6 parasites were resuspended in 1 ml of lysis buffer (10 mM HEPES pH 7.9, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 1 mM DTT, 0.5 mM PMSF, 0.65% NP-40). Nuclei were collected by centrifugation and extracted with 100 μl of extraction buffer (20 mM HEPES pH 7.9, 0.4 M NaCl, 1 mM EDTA, 1 mM EGTA, 1 mM DTT, 1 mM PMSF). After 15 min of vigorously shaking the extract was cleared by centrifugation and yields a protein concentration of 1-2 μg/ml. Five microgram of total nuclear protein was incubated with 10 fmol of double stranded, end labelled oligonucleotides for 20 min at room temperature (20 mM HEPES pH 7.9, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, 5% glycerol, 0.25 mg/ml BSA, 2 μg poly d(I/C) or as indicated; final volume: 15 μl). Binding assays were analyzed by gel electrophoresis (4% polyacrylamide, 5% glycerol, 0.5 x TBE).

RESULTS

The KAHRP gene is transcriptionally regulated

Three asexual intraerythrocytic stages have been defined for the protozoan parasite P. falciparum—the ring, trophozoite and schizont stage. The accumulation of KAHRP RNA during the intraerythrocytic cycle was examined by Northern analysis. The intraerythrocytic stages were separated by percoll/sorbitol gradient centrifugation (30). Total cellular RNA was isolated from the ring, the trophozoite, and the schizont stages of the cloned isolate FC3-A2, respectively. Unlike previously published (9), the 4.2 kb KAHRP transcript accumulates during the ring stage and only small amounts are detectable in trophozoites (Fig. 1).

Temporal changes in the KAHRP promoter activity occurring during the intraerythrocytic cycle were studied by nuclear run-on analysis (Fig. 2). For comparison the P195 gene encoding the major merozoite surface antigen (12), the CS gene encoding the circumsorozoite antigen expressed during the insect stage (13), and the ribosomal RNA genes (14) were analyzed in parallel. Nuclei were isolated from synchronized cultures of FC3-A2 at 18 hrs (ring stage), 30 hrs (trophozoite stage), and 41 hrs (schizont stage) after infection. Preformed transcriptional complexes were allowed to form in the presence of labelled UTP. The labelled nascent RNA was used as a probe for DNA fragments specific for the KAHRP, the P195, the CS, and the
rRNA genes respectively. The KAHRP gene is transcriptionally active only during the ring stage (Fig. 2). The correlation between its promoter activity and RNA accumulation indicates that the KAHRP gene is transcriptionally regulated. In contrast to the erythrocytic stage specific regulation of the KAHRP gene, the P105 gene and the rRNA genes are constitutively transcribed during these stages (Fig. 2). No transcription was observed for the CS gene during the erythrocytic cycle, consistent with its specificity for the insect stage. Transcription of the KAHRP gene and the P105 gene is sensitive to α-amanitin (Fig. 2), indicating that these genes are transcribed by an α-amanitin-sensitive RNA polymerase.

**Structural organization of the KAHRP locus**

Genomic sequences were obtained for the KAHRP gene in order to analyze the structural basis for the stage specific regulation observed for this gene. Several overlapping genomic clones spanning the entire locus were obtained and sequenced (Fig. 3). The RNA initiation site of the KAHRP gene was determined by primer extension and RNase protection assays (Fig. 4). The primer extended product was recovered from the gel and amplified by anchored PCR. Both primer extension and RNase protection experiments map the start point of transcription to a single site, 849 bp upstream of the KAHRP initiation codon (Fig. 4). The sequence 5' of the RNA initiation site was examined for homologies to the binding sites of known eukaryotic transcription factors. None were found. However, a novel sequence element with two-fold symmetry was observed 160 bp upstream of the RNA start site (position 1562 to 1573 in Fig. 3B).

**Nuclear proteins interact stage-specifically with a sequence motif**

To determine whether this motif is a binding site for nuclear proteins, gel retardation assays were performed. Oligonucleotides

**Fig. 2. Stage-specific transcriptional activity of the KAHRP gene.** The transcriptional activity of the KAHRP gene during the intraerythrocytic cycle was examined by stage-specific nuclear run on analysis. Nuclei were isolated from ring, trophozoite, and schizont S. pylonya parasites, respectively. Nascent, labelled RNA was used as a probe to DNA fragments containing the ribosomal RNA genes. Coding sequences of the KAHRP and the P105 (2) and the CS (1) genes respectively, indicated. Prior to transcription, nuclei prepared from ring stage parasites were incubated with 100 μg ml-1 α-amanitin for 15 min on ice.

---

**Fig. 3. Genomic organization and sequence KAHRP upstream sequence.** A: Genomic organization and clones. The KAHRP reading frame is indicated by rectangles. The RNA initiation site is indicated by an arrow head. The genomic clone 9034 contains an internal deletion as indicated by a triangle. B: Sequence of the KAHRP upstream region. The RNA initiation site is indicated by an arrow head. A sequence motif with two-fold symmetry is indicated by arrows. The box indicates the sequence element that is used as gel retardation assays (Fig. 5). Position 2215 EcoRI sites correspond to position 248 of the published KAHRP coding sequence (GenBank accession number J02972).
containing the KAHRP sequence element were incubated with nuclear extracts prepared from asynchronously growing erythrocytic cultures of the parasite isolate FCR3-A2. Complex formation was analyzed by gel electrophoresis. Three complexes were observed (Fig. 5A). These complexes were found to be stable in the presence of high concentrations of non-specific competitor DNA. Further, neither pUC18 DNA (1 µg) nor a DNA fragment containing the GBP130 upstream sequence could compete the formation of these complexes (Fig. 5A). By contrast, complex formation is not observed in the presence of unlabelled KAHRP oligonucleotides (100 ng) or of a DNA fragment containing the KAHRP upstream region (1 µg). Extracts prepared from uninfected red blood cells do not interact with the KAHRP sequence element. To determine the stage specificity of these complexes, the KAHRP oligonucleotide was incubated with extracts prepared from ring and schizont stage parasites, respectively (Fig. 5B). Three complexes, two major and one minor, were observed with nuclear extracts prepared from the ring stage. In contrast, the KAHRP sequence element formed one complex with nuclear extracts from the schizont stage (Fig. 5B). These data indicate that the KAHRP sequence element interacts with nuclear proteins in a stage specific manner.

DISCUSSION

The protozoan parasite *Plasmodium falciparum* has a complex life cycle alternating between a vertebrate and an invertebrate host. During the life cycle gene expression is regulated as indicated by distinct patterns of RNA (9, 15, 16, 17) and protein accumulation (10, 11, 18, 19, 20, 21). The mechanisms of gene regulation are not

![Diagram](image-url)

Fig. 4. Mapping the RNA initiation site of the KAHRP gene. RNase protection analysis. A single stranded, radiolabelled RNA probe (position 1419 to 2215 in Fig. 3B) complementary to the KAHRP RNA was generated and hybridized to 1 µg of poly A+ ring stage RNA derived from the parasite clone FCR3-A2. After RNase digestion products were analysed by gel electrophoresis. Controls using poly A+ RNA prepared from the mutant KAHRP containing the KAHRP upstream region and yeast were analysed in parallel. A size standard is indicated. Primer extension: A end labelled primer (corresponding to position 1809 to 1850 in Fig. 3B) was hybridized to 5 µg of poly A+ ring stage RNA. Extension products were analyzed by gel electrophoresis and compared to a sequencing reaction of genomic DNA using the same primer. Additional primer extended products evident on the autoradiogram were found to be caused by pausing of reverse transcriptase due to the AT richness of the sequence. The primer extended product was recovered from the gel, amplified by anchored PCR technology (35), cloned and sequenced.

![Diagram](image-url)

Fig. 5. Stage specific interaction of the KAHRP sequence element with nuclear extracts. A. 10 ng of double stranded, end labelled oligonucleotides containing the KAHRP sequence element were incubated with 5 µg total nuclear protein derived from asynchronously growing parasitic cultures. The sequence of the oligonucleotide is shown at the bottom. The amount of poly d[1419].C added to the binding assays is indicated. One µg of poly d[1419].C equals about 5 pmol of non specific binding sites. For cross competition experiments 100 ng t 5 pmol of unlabelled KAHRP oligonucleotides or 1 µg (6.6 nmol of non specific binding sites) of pUC18 DNA or of DNA fragments containing the upstream regions of the KAHRP and GBP30 (25) genes, respectively, were tested for their ability to compete. A control experiment was performed using extracts from uninfected erythrocytes (lane marked RBC extract). B. Nuclear extracts were prepared from ring (R), and schizont (S) stage parasites, respectively. The extracts were incubated with 1 µl labelled KAHRP sequence element in the presence of 2 µg of poly d[1419].C. Where indicated 100 ng of unlabelled KAHRP oligonucleotides were added to the binding assay. Complexes formed are indicated by arrows.
well understood for this parasite, mainly because functional assays for the study of promoter activity, i.e., transcription or a reconstituted in vitro transcription system, are not yet available. In addition, structural data are difficult to obtain, since non-coding plasmoidal DNA are unstable in E. coli (22), which presumably results from the unusually high A/T content of $>90\%$ (23, 24).

Consequently, structural elements involved in the developmental expression of P. falciparum genes have not been defined. Knowledge of these elements, however, is a prerequisite for the development of a transcription system for P. falciparum and will help in our understanding of the mechanisms of gene regulation in this parasite. In this study, we have identified a structural element in the upstream region of the KAHRP gene that, in correlation with the transcriptional activity of this gene during the intraerythrocytic cycle, binds to parasite derived nuclear factors in a stage specific manner.

A single RNA initiation site has been found for the KAHRP gene as determined by both primer extension and RNA protection experiments. A unique transcription start site has also been noted for the erythrocytic stage gene GPB130 of P. falciparum (25). In contrast, multiple RNA initiation sites have been reported for the insect stage circumsporozoite gene of the simian parasite P. knowlesi (26), the erythrocytic stage major merozoite surface antigen gene P195 from P. falciparum (27), and the P.240 gene from the rodent parasite P. yoelii (28). The frequent pausing of reverse transcriptase in A/T rich regions can easily be misinterpreted as multiple initiation sites. Hence, it has yet to be determined whether plasmoidal genes can have either single or multiple RNA initiation sites. A comparison of genomic and cDNA sequences indicates that the KAHRP gene is continuously transcribed. However, post-transcriptional processing of the KAHRP transcript occurs through 3'-splicing and polyadenylation.

A comparison of transcriptional activity as determined by nuclear run-on analysis and RNA accumulation indicates that the KAHRP gene is transcriptionally regulated during the erythrocytic cycle and transcribed only during the ring stage. Like the KAHRP gene, the GPB130 (25) and the HRP II (data not shown) genes are also transcriptionally regulated, while the P195 gene is constitutively transcribed in the erythrocytic stages (Fig. 2).

Transcription of the KAHRP gene is sensitive to the RNA polymerase inhibitor a-amanitin. Eukaryotic genes transcribed by an a-amanitin sensitive RNA polymerase frequently contain regulatory signals immediately upstream of the RNA initiation site. A comparison of the upstream regions of several plasmoidal genes including the GPB130 and the CS genes has revealed a common sequence element that is homologous to the core region of the SV40 enhancer (25, 26). This SV40 enhancer-like element is not present in the KAHRP upstream region, nor were there any homologues to other known binding sites of eukaryotic transcription factors. However, a palindromic sequence element was identified 160 bp upstream of the RNA initiation site of the KAHRP gene. This element was recognized in a specific fashion by parasite derived nuclear extracts. Different complexes were formed with this element depending on the developmental stage of the parasite. Three distinct complexes of different mobility were observed in the ring stage during which the KAHRP promoter is transcriptionally active. During the schizont stage, when the KAHRP promoter is silent, a single complex is formed.

It is tempting to conclude that the KAHRP palindromic motif is a promoter component which mediates the developmental expression of the KAHRP gene by interacting with transcriptional factors in a stage specific manner. It remains to be determined whether the protein-DNA complexes observed define multiple protein binding sites on the KAHRP sequence element or whether a single DNA binding protein is post-transcriptionally modified or binds additional co-factors. It is also possible that this sequence element is recognized both by positive regulatory factors during the ring stage and by negative regulatory factors during the schizont stage. Until a functional promoter assay is developed, these conclusions remain speculative.

ACKNOWLEDGEMENTS

This work was supported by the U.S. Army grant DAMD 17-89-Z-0003 and the World Health Organization. J. V. R. is a scholar in molecular parasitology supported by the Burroughs-Wellcome Fund. M. L. is supported by fellowships from the Deutscher Akademischer Austauschdienst and from the World Health Organization. Technical assistance was provided by Carter Ralph and Mark Samuel. Secretarial support was provided by Cynthia Ritter.

REFERENCES


