AD_____ (2)

# MOLECULAR GENETIC ANALYSIS OF PARASITE SURVIVAL IN P. FALCIPARUM MALARIA

ANNUAL REPORT

JEFFREY V. RAVETCH



MARCH 15, 1991

Supported by

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

Contract No. DAMD17-89-Z-9003

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

# REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188



| 1. AGENCY USE ONLY (Leave blank) | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | March 15, 1991 | Annual 1 Nov 89 - 30 Nov 90 |

**4. TITLE AND SUBTITLE**
Molecular Genetic Analysis of Parasite Survival in *P. falciparum* Malaria

**5. FUNDING NUMBERS**
DAMD17-89-Z-9003
61102A
3M161102BS13 AF
DA315681

**6. AUTHOR(S)**
Jeffrey V. Ravetch

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Sloan-Kettering Institute
Memorial Sloan-Kettering Cancer Center
1275 York Avenue
New York, New York 10021

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
U.S. Army Medical Research and Development Command
Fort Detrick
Frederick, Maryland 21702-5012

**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**

**12a. DISTRIBUTION/AVAILABILITY STATEMENT**
Approved for public release; distribution unlimited

**12b. DISTRIBUTION CODE**

**13. ABSTRACT** *(Maximum 200 words)*

The human malaria parasite evades the protective mechanisms of its host through a complex variety of strategies. Extensive genetic variation is thought to contribute to the mechanism by which the parasite survives within its vertebrate host. Recent studies have focussed on the mechanism of genetic variability of the parasite. Pronounced chromosomal size variations are observed between diffe ent geographical isolates of the parasite and during mitotic growth of the p asite in culture. Several of these chromosomal polymorphisms have been characcerized and found to be the result of chromosome breakage followed by the healing of these broken ends by the addition of telomere repeats, resulting in large distal deletions and truncated chromosomes. Only the chromosomal gragments associated with centromere containing elements are mitotically stable and retained. Further insight into the molecular mechanism of this process was obtained by the analysis of the RESA inversion/breakage on chromosome 1. In at least one case, this process resulted in a transcriptional

**14. SUBJECT TERMS**
Malaria, Vaccine, Molecular biology, Merozoite, Erythrocyte, Recombinant DNA, RA 1

**15. NUMBER OF PAGES**

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTR. |
|---|---|---|---|
| Unclassified | Unclassified | Unclassified | Unlimited |

19. Abstract (continued)

initiation site being brought into the vicinity of a telomere. Transcription from this site was no longer observed in its new chromosomal position. Studies were pursued to determine how the position of a gene along the P. falciparum chromosome can influence its expression. Detailed analysis of transcriptional start sites for the GBP130 and KAHRP genes was completed. The sequence elements involved in transcriptional initiation were identified.

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 Army endorsement or approval of the products or services of these organizations.

__✓__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 Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

_____For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

__✓__In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institute of Health.

Jeff V Rouhl 7/24/91

PI Signature Date

## TABLE OF CONTENTS

## BODY OF REPORT

### BACKGROUND AND SIGNIFICANCE

The genome of Plasmodium falciparum, the protozoan parasite responsible for the most severe form of human malaria, is extremely flexible. It displays polymorphisms on the order of fifty to hundreds of kilobase pairs between genetically equivalent chromosomes from different parasite isolates (1, 2, 3). These chromosome length polymorphisms can involve deletions in which structural genes are lost (4, 2, 5). One well-characterized mechanism by which this class of polymorphisms is generated is the introduction of double-stranded breaks in the DNA, followed by the enzymatic addition of telomere repeats to the free 3' ends, thereby stabilizing the foreshortened chromosome fragment to mitosis (5). A conserved sequence element has been identified at the breakpoint of multiple, independent events, suggesting a specificity to either the breakage or the healing reaction (5). This specificity has been further characterized by the detailed analysis of a subtelomeric rearrangement for chromosome 1 in the ring-infected erythrocyte surface antigen (RESA) gene.

DNA rearrangements in P. falciparum frequently result in the loss of transcriptional activity of the effected gene. This is due either to deletion of putative promoter elements, inversion of these elements or juxtaposition of promoter sequences to telomere regions. To define the mechanism involved in these transcriptionally inactive states, characterization of transcription units for the parasites have been pursued.

The human malaria parasite, Plasmodium falciparum, has a complex life-cycle, in which the protozoan parasite alternates between man and mosquito. Within the human host the parasite multiplies asexually within the erythrocyte, resulting in repeated cycles of invasion and lysis. At least three distinct morphological stages have been defined during the intraerythrocytic stage of parasite development, the ring, trophozoite and schizont. These intracellular stages can be readily separated in P. falciparum cultures (6, 7, 8) and display distinct patterns of protein expression (9, 10). Steady-state levels of RNA accumulation for several erythrocytic stage genes have been shown to be stage-specific (11, 2, 9). However, the molecular mechanisms which regulate this expression remain obscure.

Transcription in Plasmodium in general, and P. falciparum specifically, is largely uncharacterized. This is due to the absence of a functional assay to determine promoter activity. Transfection of these protozoans or a reconstituted in vitro transcription system are not available. Putative promoter regions have been investigated for only two genes, the insect stage circumsporozoite gene of the simian parasite P. knowlesi (12) and a blood stage antigen, Py230, from the rodent parasite P. yoelli (13). For P. falciparum neither upstream regulatory regions nor

transcriptional start sites have been reported to date. This is due, in part, to the difficulty in maintaining DNA fragments derived from P. falciparum in E. coli hosts, due to the extreme A+T richness of this organism's genome. Thus, the elements which constitute a promoter in this protozoan have not been defined, nor the cis-acting elements which regulate stage-specific expression.

To define these mechanisms in this important human pathogen and to establish the structural components needed for a transfection system, we have characterized RNA transcription in a variety of blood-stae genes and have defined putative promoter regions for three genes. A short intergenic region has been found between two genes transcribed during the blood stage in which the termination of one transcript and initiation of a second transcript are defined. Comparison of the promoter sequences for several genes has revealed common structural patterns with elements of both eukaryotic and prokaryotic promoters. These elements can be specifically bound by parasite derived nuclear proteins. For one element stage specific binding is demonstrated suggesting that these elements may be involved in gene regulation.

## PROGRESS REPORT

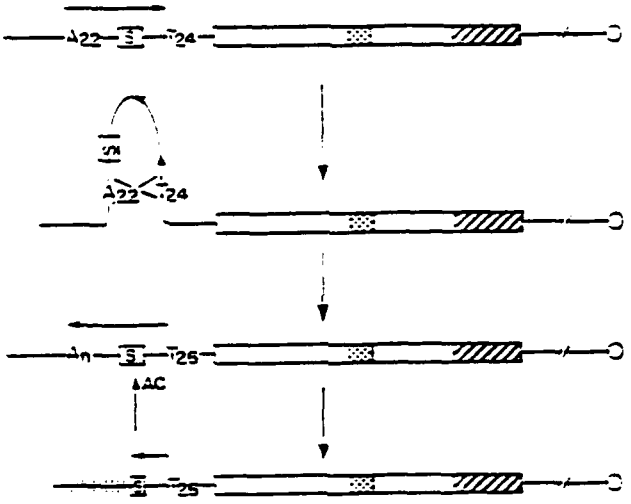
**Mechanism of RESA inactivation** (Pologe, de Bruin and Ravetch submitted 12/89; in press 2/90).

*RESA is a 155,000-$M_r$ peptide of unknown function* which is deposited onto the erythrocyte surface by the invading merozoite during asexual development (14) and which accumulates in the erythrocyte cytoplasm surrounding the gametocyte at the onset of sexual development (15). The structure of the RESA gene in a RESA$^-$ isolate has been shown to involve both an inversion of the signal-encoding exon and a large deletion. This inversion is mediated by homologous recombination between homopolymeric A and T sequences in the 5' untranslated and intron sequences, respectively. The deletion includes DNA sequences from the middle of the signal exon 5' to the end of the chromosome. The resulting truncated gene is followed by an abrupt transition to the P. falciparum telomere repeat sequence.

In order to determine the mechanism and generality of the RESA gene deletion and chromosome 1 rearrangement, restriction mapping was performed on DNAs from a number of P. falciparum isolates.

A 7.8-kilobase fragment was detected in the DNA from RESA$^-$ isolate FC27, as well as in DNAs from clone D10 and nonclonal isolate FCR3. However, five FCR-3 derived clones displayed a smaller HincII band and migrated as heterogeneous-sized DNA fragments. BAL 31 susceptibility of the RESA gene in total genomic DNA from one of these clones demonstrated that the gene had assumed a telomeric location in these parasites (E96; data not shown). Cloning and sequencing of the D3 mutant revealed an organization in which a 5' sequence was inverted and an abrupt transition occurred to telomere repeats, similar to events described for the

KAHRP on chromosome 2 and HRPII on chromosome 8. The breakpoint occurs following a CA dinucleotide sequence, as has been found immediately preceding the transition to telomere repeats in every gene rearrangement to a telomere in P. falciparum described to date (5). This suggests that a common sequence element is necessary for some step in the pathway to a viable rearrangement. This observation leads us to propose a model (as shown below) in which an inversion, which is mediated by base pairing and a crossover event between the homopolymeric A and T sequences that flank the inverted sequences, precedes the deletion event. Subsequent DNA breakage and telomere addition generated the observed chromosome 1 structure.




Proposed mechanism for generating the RESA⁻ structure. The RESA gene is located on chromosome 1 and is oriented 5' to 3', telomere to centromere. Symbols: ○, centromere; ∥, indeterminant number of kilobase pairs; □, translated sequences of the RESA gene with the nontranscribed poly(dA) and poly (dT) sequences 5' of the signal encoding sequences (S) and second exon, respectively; ▭ and ▨, sequences encoding repeated amino acids (the direction of transcription is 5' to 3' from left to right); — above the signal exon, sequence that is inverted in the pathway; ———, telomere repeats. The conserved CA at the breakpoint is shown in line 3 oriented 5' to 3', centromere to telomere.

This model predicts the existence of an intermediate in which the signal exon and the intron sequences are inverted. To identify this intermediate, the polymerase chain reaction (PCR) was used with oligonucleotides which would specifically target the inverted structure. Total genomic DNAs from a RESA$^-$ mutant and from a number of parasite isolates in which the RESA gene appeared unrearranged were characterized.

Oligonucleotide primers were synthesized which would amplify the following species: the wild type gene, the inverted gene, the inverted and rearranged gene by PCR of genomic DNA. Inversion was detected in the absence of rearrangement; however, rearrangement was seen only for inverted sequences. These studies indicate that an inversion intermediate can be detected and is likely to precede the deleted, healed form, as predicted in the model proposed above. A previous model (16) for generating the RESA$^-$ gene structure which proposed an unresolved recombination event with nuclease digestion and healing instead of an intermediate is therefore incorrect. The inversion appears to be a required first step in this DNA rearrangement. If specificity resided in the cleavage reaction, then healing could generate a RESA$^-$ telomeric gene in the absence of inversion. We have never detected a RESA$^-$ telomeric gene with the signal exon in an uninverted orientation. The inversion rearranges the conserved CA dinucleotide to a centromere-proximal position. Breakage and healing of the inverted sequence then result in a mitotically stable, truncated chromosome. These results indicate a sequence specificity for the healing reaction, which always includes a CA dinucleotide. Whether similar sequence specificity is required for the breakage event is still undetermined.

## REFERENCES


1. Kemp, D.J., Thompson, J.K., Walliker, D., and Corcoran, L.M. 1987. Molecular karyotypes of Plasmodium falciparum: conserved linkage groups and expendible histidine-rich protein genes. Proc. Natl. Acad. Sci. USA 84:7672-7676.

2. Pologe, L.G., and Ravetch, J.V. 1986. A chromosome rearrangement in a P. falciparum histidine rich protein gene is associated with the knobless phenotype. Nature (London) 322:474-477.

3. van der Ploeg, L.H.T., Smits, M., Ponnudurai, T., Vermeulen, A., Meuwissen, J.H.E., and Langsley, G. 1985. Chromosome-sized DNA molecules of Plasmodium falciparum. Science 229:658-661.

4. Corcoran, L.M., Forsyth, K.P., Bianco, A.E., Brown, G.V., and Kemp, D.J. 1986. Chromosome size polymorphisms in Plasmodium falciparum can involve deletions and are frequent in natural populations. Cell 44:87-95.

5. Pologe, L.G., and Ravetch, J.V. 1988. Large deletions result from breakage and healing of P. falciparum chromosomes. Cell 55:869-874.

6. Lambros, C., and Vanderberg, J.P. 1979. Synchronization of Plasmodium falciparum erythrocytic stages in culture. J. Parasitol. 65:418-420.

7. Pasvol, G., Wilson, R.J.M., Smalley, M.E., and Brown, J. 1978. Separation of viable schizont-infected red blood cells of Plasmodium falciparum from human blood. Ann. Trop. Med. Parasitol. 72:87-88.

8. Trager, W.T., and Janse, J.B. 1976. Human malaria parasites in continuous culture. Science 193:673-675.

9. Weber, J.L., 1988. Molecular biology of malaria parasites. Exp. Parasitol. 66:143-170.

10. Wesseling, J.G., Snijders, P.J.F., Someren van, P., Jansen, J., Smits, M.A., and Schoenmakers, J.G.G. 1989. Stage-specific expression and genomic organization of the actin genes of the malaria parasite Plasmodium falciparum. Mol. Biochem. Parasitol. 35:167-176.

11. Kochan, J. Perkins, M., and Ravetch, J.V. 1986. A tandemly repeated sequence determines the binding domain for an erythrocyte receptor binding protein of P. falciparum. Cell 44:689-696.

12. Ruiz i Altaba, A., Ozaki, L.S., Gwadz, R.W., and Godson, G.N. 1987. Organization and expression of the Plasmodium knowlesi

circumsporozoite antigen gene. Mol. Biochem. Parasitol. 23:233-245.

13. Lewis, P.A. 1990. Sequence analysis upstream of the gene encoding the precursor to the major merozoite surface antigens of Plasmodium yoelli. Mol. Biochem. Parasitol. 39:285-288.

14. Brown, G.V., Culvenor, J.G., Crewther, P.E., Bianco, A.E., Coppel, R.L., Saint, R.B., Stahl, H.-D., Kemp, D.J., and Anders, R.F. 1985. Localization of the ring-infected erythrocyte surface antigen (RESA) of Plasmodium falciparum in merozoites and ring-infected erythrocytes. J. Exp. Med. 162:774-779.

15. Uni, S., Masuda, A., Stewart, M.J., Igarashi, I., Nussenzweig, R. and Aikawa, M. 1987. Ultrastructural localization of 150/130 Kd antigen in sexual and asexual blood stages of Plasmodium falciparum-infected human erythrocytes. Am. J. Trop. Med. Hyg. 36:481-488.

16. Cappai, R., van Schravendijk, M.-R., Anders, R.F., Peterson, M.G., Thomas, L.M., Cowman, A.F., and Kemp, D.J. 1989. Expression f the RESA gene in Plasmodium falciparum isolate FCR3 is prevented by a subtelomeric deletion. Mol. Cell Biol. 9:3584-3587.

## BIBLIOGRAPHY 1989-1990


1. Pologe, L.G., de Bruin, D. and Ravetch, J.V. (1990). A and T homoplymeric stretches mediate a DNA inversion in Plasmodium falciparum which results in loss of gene expression. Mol. Cell. Biol. 10:3243-3246.

2. Lanzer, M. and Ravetch, J.V. Conserved DNA sequences associated with P. falciparum transcription units are the sites of specific protein interactions. Submitted, 1990.

**Review**

3. Ravetch, J.V. (1989). Chromosomal polymorphisms and gene expression in Plasmodium falciparum. Exp. Parasitol. 68:121-125.

# APPENDIX

Conserved DNA sequences associated with P. falciparum transcription units are the sites of specific protein interactions (Lanzer and Ravetch, submitted 1990)

Stage-specific protein expression occurs during the intraerythrocytic cycle of the human malaria parasite and is responsible for the unique morphological and pathogenic properties of this protozoan. We have characterized the transcription of several erythocytic-stage genes to determine the molecular basis for this specificity. In most cases, promoter activity, as measured by nuclear run-on assays, parallels RNA and protein accumulation. The putative promoters for three erythrocytic stage genes have been isolated and characterized. An intergenic region has been identified between two tightly linked transcription units, defining the minimal and necessary sequences for a transcriptional termination and initiation signal in Plasmodium. Sequence comparison of the promoter regions of multiple genes has revealed two classes of sequence elements: enhancer-like elements and interspersed, G/C rich repetitive elements. A functional role for these elements is suggested in gel-retardation assays which demonstrate that these elements are the binding sites for parasite derived nuclear proteins.

## Introduction

The human malaria parasite, Plasmodium falciparum, has a complex life-cycle, in which the protozoan parasite alternates between man and mosquito. Within the human host the parasite multiplies asexually within the erythrocyte, resulting in repeated cycles of invasion and lysis. At least three distinct morphological stages have been defined during the intraerythrocytic stage of parasite development, the ring, trophozoite and schizont. These intracellular stages can be readily separated in P. falciparum cultures (8,12,17) and display distinct patterns of protein expression (20,21 ). Steady-state levels of RNA accumulation for several erythrocytic stage genes have been shown to be stage-specific (7,13,20). However, the molecular mechanisms which regulate this expression remain obscure.

Transcription in Plasmodium in general, and P. falciparum specifically, is largely uncharacterized. This is due, in part, to the difficulty in maintaining DNA fragments derived from P. falciparum in E. coli hosts, due to the extreme A+T richness of this organism's genome. Thus, transcriptional start sites have been investigated for only two genes, the insect stage circumsporozoite gene of the simian parasite P. knowlesi and a blood stage antigen, Py230, from the rodent parasite P. yoelii. Multiple initiation sites were suggested for each gene from those studies (10,15). Thus, the elements which constitute a promoter

in this protozoan have not been defined, nor the cis-acting elements which regulate stage-specific expression.

To define these mechanisms in this important human pathogen, we have characterized RNA transcription in a variety of blood-stage genes and have defined the structural components of putative promoters for three genes. A short intergenic region has been found between two genes transcribed during the blood stage in which the termination of one transcript and initiation of a second transcript are defined. Comparison of the promoter sequences for several genes has revealed common structural patterns with elements of both eukaryotic and prokaryotic promoters. These elements can be specifically bound by parasite derived nuclear proteins.

## Results

### Stage-specific gene expression is transcriptionally regulated

Temporal changes of transcriptional activity during the asexual blood cycle were studied by nuclear run-on analysis of six previously described genes. Four blood-stage specific genes were studied, encoding the major merozoite antigen (P195), the glycophorin binding protein (GBP-130), the knob associated histidine-rich protein (KAHRP) and the histidine-rich protein II (HRP II). A gene expressed only during the insect stage, the circumsporozoite antigen (CS), and genes which are constitutively expressed, the ribosomal genes, were analysed in parallel for comparison.

Nuclei were prepared from synchronously growing cultures at 18hr (ring stage), 30hr (trophozoite stage) and 41hr (schizont stage) after infection. The nuclei were transcribed *in vitro* in the presence of radiolabeled UTP. The radiolabeled RNA was isolated and hybridized with DNA samples. Stage-specific transcription is observed for the four blood stage genes, in contrast to the constitutive expression of the ribosomal genes (Fig. 1A). The KAHRP gene is transcriptionally-active only during the ring stage, while the GBP-130 gene is transcribed in trophozoites and in schizonts. Similarly, the HRP II gene is transcriptionally active during the ring and trophozoite stage and silent during the schizont stage. In contrast to the tight

stage specific regulation of the these genes, the P195 gene is transcribed in all asexual blood stages. Its promoter activity, however, changes from 60-70ppm in rings and schizonts to 480ppm observed in the trophozoite stage (Table 1). This promoter activity is more than an order of magnitude greater than the other blood stage genes, exceeding the transcriptional activity of the ribosomal gene promoter (Table 1). No transcription was observed for the CS gene during the blood stage (not shown), consistent with its specificity for the insect stage.

To determine if post-transcriptional events are significant in the regulation of the expression of these blood stage genes, we compared the transcriptional activity to the accumulation of total cellular RNA. RNA was isolated from synchronized cultures at the same time points after infection as above and analysed by Northern analysis. With the exception of the KAHRP gene, promoter activity and RNA accummulation correlate well (Fig. 1B). Steady state RNA accumulation for the KAHRP gene is detectable at the ring and trophozoite stage. Its promoter, however, is active only during the ring stage (Fig. 1A) indicating that the stability of the KAHRP RNA contributes to its accumulation during the trophozoite stage. Transcription of these blood stage genes is sensitive to α-amanitin, while ribosomal gene transcription is resistant (Fig. 1C) The KAHRP-, the GBP-130-, the ERPII, and the P195 gene are therefore transcribed by an alpha amanitin-sensitive, Pol II-like RNA polymerase.

Isolation of erythrocytic stage promoter-containing sequences

The structural basis for the developmental changes of promoter activity was analysed by cloning the upstream regions of three genes, the GBP-130, the KAHRP, and the P195. Two genomic libraries of Plasmodium falciparum (strain A2), a pUC9 plasmid and a lambda gt11 library, were screened with DNA probes corresponding to sequences at the 5' end of the published sequences. As has been observed previously (7), many of the clones isolated contained large internal deletions which occured between long runs of alterating A and T nucleotides. Undeleted sequences were obtained by the propagation of these DNA fragments in the recombination defective host SURE. Existing deletions were filled by PCR amplification of genomic DNA, cloned into pUC18 and propagated in SURE. The integrity of all clones and sequences determined were confirmed by correlation of the predicted restriction maps of the cloned fragments with genomic DNA and by PCR amplification of genomic DNA yielding fragments of the predicted size (data not shown).

Two blood stage transcription units are tightly linked

To obtain sequences upstream of the published GBP-130 sequence, 300,000 plaques of a lambda gt11 genomic library of Plasmodium falciparum were screened with the cDNA clone 2374 (7)

(Figure 2A). One clone was isolated (8771). In addition, one clone (2044) previously published was obtained from a plasmid library. Since the the clone 2044 contains a large internal deletion which is only partialy filled by the clone 8771, the remaining gap was closed by PCR amplification of genomic DNA yielding the clone A420. The compiled sequence is shown in figure 2B. 5' of the coding region of the GBP-130 gene the A/T content increases to over 85%. The first 720bp of this sequence contain features characteristic of a plasmodial gene. A continuous open reading frame of 93 amino acids is found, followed by an intron of 201 bp, flanked by consensus splice donor and acceptor sequences. The final exon encodes 26 amino acids and a short 3' untranslated sequence containing a polyadenylation site. To confirm the presence of this tightly linked transcription unit, Northern analysis was performed with total cellular RNA isolated from asynchronous growing blood cultures and hybridized with a probe derived from this putative transcription unit (Xba/NcoI fragment) (Fig. 3). A novel RNA species of 3.8kb is seen (Fig. 3A). Hybridization of the same RNA with a fragment containing transcribed sequences of the GBP130 gene (2734) reveals the GBP-130 transcript of 6.6kb (Fig. 3B) as do cDNA probes specific to the GBP130 gene. To determine the polarity of this transcript, a single stranded probe was generated by SP6 RNA polymerase upon cloning the XbaI/HindIII fragment of 2044 into pGEM3. As seen in Fig. 3A both the 3.8 kb RNA and the GBP-130 RNA are derived from the same DNA strand, confirming that both genes are transcribed

in the same direction. cDNA clones have been isolated from a blood stage library spanning the entire 3.8kb (Fig. 2A). No sequence homology to any known gene has been detected for this open reading frame.

Transcription across this intergenic region is discontinuous, as revealed by nuclear run-on analysis. Nuclei were labelled during the schizont stage (the transcriptionally active stage of the GBP-130 gene) and the radiolabelled RNA used as a probe for DNA fragments derived from the intergenic region (Fig. 3B). Less radiolabelled RNA hybridized to a DNA fragment taken from the intergenic region than to probes containing sequences from the coding region of the 3.8 and the GBP130 gene (Fig. 3B). The stage-specificity of transcription for the 3.8kb transcript has not been determined, although it is weakly transcribed during the schizont stage. The 3.8kb transcript is terminated with an efficiency of approximately 70% as calculated from the ratio of bound radioactivity (upon correcting for the T content of the the various DNA fragments). The sequence of the terminator is likely to be AAAATAAAA followed after 100 nucleotides by a run of 30 As, as has been observed in other lower eukaryotes (6). The initiation site for the GBP-130 gene was determined by S1 mapping and primer extension (Fig. 3C).

These data indicate that the GBP-130 gene is closely linked at its 5' end to another gene which is also transcribed during the blood stages. The sequences between both transcribed regions must contain the minimal elements which signal transcription

termination and initiation in P. falciparum blood stage genes.

Stuctural analysis of blood-stage promoter regions

The sequences immediately 5' of the initiation site of the GBP130 gene contain a variety of structural features indicative of a putative promoter (Fig. 2B). While the extraordinary A+T content of this region precludes the assignment of a TATAA sequence, an element at positions -114 to -141 contains homology to the SV40 enhancer sequence (Table 2). No further homologies to known protein binding sites were found. Two additional, prominent features of the intergenic region of the GBP130 gene are indicated in Fig. 2B. Between position -315 and -948 the GBP130 intergenic region contains a duplication of 305bp, indicated by the enclosed sequences. Within the duplicated region and downstream of it are found short, G/C rich reiterated sequences. The sequence TGTGTAC is followed by TY(T/A)CCCY(T/G)T. Each sequence is repeated five times.

Functional characterization of these sequences either by transfection into P. falciparum or in vitro transcription are not possible at the present time. However, if these elements represent functional components of the transcriptional signals for blood stage genes, we would expect to find this structural organization conserved for other genes regulated in a similar fashion. In order to compare the structural organization of the GBP130 intergenic region with other transcriptional units the

upstream regions of the KAHRP (Fig. 4) and of the P195 (Fig. 5) gene were cloned. The start point of transcription of both genes were determined by primer extension and by S1 mapping (data not shown).

As summarized in Fig. 6, the upstream regions of the KAHRP and P195 contain the same structural organization as the GBP130 intergenic region: enhancer-like sequences as well as short, G/C rich repeated elements. Within this conserved organization, sequence elements are observed which are retained among all three genes, as well as sequences which appear to be gene specific. For example, the conserved sequence TRCATGTA is repeated 3 times in the P195 and KAHRP upstream region and is found in a single copy in the GBP130 intergenic region. In contrast, the repeated element TATACATATGTG appears to be unique to the P195 gene. Neither the P195 nor the KAHRP gene contain the SV40 enhancer-like sequence found in the GBP130. In its place is found the sequence TAGTGYA(C/G)TAA in two copies in the P195 and in single copy in the KAHRP upstream sequence (Table 2). The GBP130 and P195 harbor repetitive elements not only in their upstream regions but also within the first hundred basepairs of the 5' transcribed region.

Multiple geographic isolates of P. falciparum have been described which demonstrate significant genetic variation. Some of the major protein products of this parasite have been found to be highly polymorphic, and the genes which encode them show minimal sequence conservation in coding sequences and

divergent sequences in non-coding regions. If the repetitive element associated with these transcription units have a biological function then they might be expected to be conserved among various strains and isolates of Plasmodium falciparum. This hypothesis was tested for the two elements TGTGTAC and TATACATATGTG. These sequences contain the recognition sites for restriction endonucleases, RsaI and Nde I, respectively. Genomic DNA of 16 different strains and clones representing 8 independent isolates were cut by Rsa I and NdeI respectively and analysed by Southern analysis. The TGTGTAC element was found to be conserved in the GBP130 and KAHRP genes in all strains tested, while the TATACATATGTG element, found only in the P195 upstream sequence was present in that gene for all the isolates tested, despite the significant polymorphism which has been described for this gene in different isolates (5).

Nuclear extracts interact with conserved sequence elements

To determine whether the conserved sequence elements are the target sites of specific protein interactions, gel retardation assays were performed. Nuclear extracts were prepared from asynchronously growing blood cultures of Plasmodium falciparum, incubated with double stranded oligonucleotides and analysed by gel electrophoresis. The double stranded oligonucleotides used contain the conserved enhancer-like sequence elements either AACTGCATGTAGTGTAGTAA (oligo 1) which is found in the upstream

region of the KAHRP and of the P195 gene or AAATGTAAGCAGAAAAGGAATGGTGTGTTAACTTAT (oligo 2) which occurs in the intergenic region of the GBP130 gene and which shares homology to the SV40 enhancer. Both DNA fragments form stable complexes with nuclear proteins even in the presence of 6ug of poly dI/dC (Fig. 7). No binding occurs when double stranded oligonucleotides unrelated to the conserved elements of Plasmodium falciparum were used (lane marked random oligo). Furthermore, extracts prepared from uninfected red blood cells do not bind any of the sequences tested. The specificity of the complexes observed were further analysed by competition experiments. Plasmid DNA containing either the upstream region of the KAHRP or of the P195 gene compete very efficiently for complexes formed between nuclear proteins and oligo 1, since both genes contain that element. By contrast, oligo 1 complexes could not be competed by plasmid DNA harboring the GBP130 upstream region or by an 500 fold molar excess of either oligo 2, pUC18 plasmid-DNA or of unrelated oligonucleotides. Similarly, cross-competition experiments were performed with oligo 2 complexes; only oligo 2 or plasmids containing this sequence (GBP130) could compete. Single site mutations introduced into these conserved sequence elements decrease the equilibrium binding constant several fold (not shown), further suggesting that the complexes formed are specific.

These experiments demonstrate that conserved sequence elements derived from the upstream region of Plasmodium

*falciparum* genes interact specifically with nuclear proteins.

## Discussion

The analysis of four blood stage specific P. falciparum genes has demonstrated that transcriptional regulation is an important mechanism for controlling gene expression in this parasite. An α-amanitin sensitive RNA polymerase, likely the homologue of RNA polymerase II, is responsible for the transcription of these genes. In most genes transcribed by RNA polymerase II, the upstream regions contain specific sequences which determine the transcriptional activity of the gene. To determine the structural basis for the stage-specific expression of plasmodial genes, we have characterized the upstream sequences for three of these genes.

The observation that two blood stage genes, the 3.8kb transcript and the GBP130 are transcribed in a head-to-tail arrangement, has defined a 2 kb region in which transcriptional termination and initiation occurs. Nuclear run-on analysis across this intergenic region revealed a discontinuity in transcription. Polycistronic transcripts were not observed, nor was a trans spliced transcipt generated from this region. Thus, despite the absence of a functional assay for plasmodial promoter activity, the intergenic region of these linked genes is likely to define the minimal structural elements necessary for transcriptional termination and initiation.

This tight linkage of transciptional units in P. falciparum has not been observed before, but may not represent an unusual

structural arrangement in this organism. Mapping of blood stage genes on the 14 P. falciparum chromosomes has suggested a clustering of these sequences at the ends of chromosomes, with the central regions appearing relatively barren (3). Whether this clustering of genes reflects a similar linkage of transcriptional units has yet to be determined. The functional significance of this clustering may reflect aspects of coordinate regulation of genes expressed at similar stages.

Structure of a P. falciparum termination signal

The sequence AAATAAA followed by a run of 30 A residues within 100bp has been shown to function as a polyadenylation and termination signal in Dictiostelium (6). This sequence arrangement is found at the 3' end of the gene encoding the 3.8kb transcript and has been observed in the 3' sequences of the simian malaria parasite P. knowlesi circumsporozoite gene (15). cDNA clones have been isolated for the 3.8kb transcript which terminate 24 nucleotides 5' of the AAATAAA sequence (data not shown) and transcriptional activity has not been detected 3' of the poly A sequence located at position -1790 (Fig. 3B). Other plasmodial genes have not been characterized in this region, precluding additional comparisons.

The putative P. falciparum promoter contains elements of both prokaryotic and eukaryotic promoters

The transcriptional initiation site for three blood stage genes of P. falciparum have been determined. All three genes have a single start point of transcription, in contrast to previous reports where multiple RNA initiation sites have been suggested in the mapping of the circumsporozoite gene of P. knowlesi and the Py230 antigen gene of P. yoelii. These differences are likely to result from differences in the methods used to map the initiation sites. It is likely that the CS gene and the Py230 gene also contain unique initiation sites.

The structure of three blood stage P. falciparum upstream sequences, containing putative promoter elements has revealed a common structural organization. When a rodent blood stage gene, Py230, and the insect stage gene from P. knowlesi, the CS gene, are included in this comparison, general features of plasmodial promoters are suggested. As schematically illustrated in Fig. 6, all genes contain two types of elements: an enhancer-like sequence, located with the first 500 nucleotides of the RNA initiation site, present in one or two copies and interspersed, G/C rich repetitive elements, present in multiple copies, which flank this element. In the GBP130 and CS genes, the enhancer element is homologous to the SV40 enhancer, as indicated in Table 2A. The KAHRP, the P195 and the Py230 genes share a different

type of putative enhancer-like element (Table 2B). The P195 of P. falciparum and Py230 of P. yoelii are related genes encoding homologous merozoite surface proteins in different species. As expected, there is little sequence homology between these genes outside of protein coding regions. However, in the sequences 5' of the RNA initiation site both genes share four elements including two copies of the enhancer-like elements (Fig. 6, Table 2B). This conservation of sequences suggests a functional role for these elements. No other sequence elements found in eukaryotic promoters have been identified in these upstream sequences, although the extreme A+T richness of this sequence presents the possibility of TATAA sequences in abundance. The enhancer-like elements have been shown to be the site of specific parasite deived nuclear proteins, further suggesting that they are functional components of the P. falciparum blood stage promoter.

The function of the interspersed G/C rich elements is not yet known. However, the elements TGTGATC and TATATATCATGTG are highly conserved among different strains and isolates of P. falciparum. It seems that some elements are not only conserved between strains but also occur in the putative promoter regions of different genes. The repeated sequence TGCATGCA found in the upstream region of the CS gene of P. knowlesi is very simmilar to the element TRCATGTA found repeated three times in the KAHRP and in the P195 upstream region. Another example of this conservation can be found for the interspersed elements CACCCCTC and

TY(T/A)CCCT(T/G)T which occur in five copies in the CS upstream region and in the GBP130 promoter region, respectively.

Interspersed repeated elements are a characteristic of cooperative protein-DNA interactions, as seen in the repression of procaryotic operons (deo or lac operon). In both of these operons three operators located within the promoter region or close to the start point of transcription and several hundred base pairs downstream within the coding sequence interact cooperatively with repressor molecules. As a result of this interactions an active RNA polymerase promoter complex cannot be formed. We can speculate that since the G/C rich elements found in putative plasmodial promoters fail to bind nuclear proteins derived from blood stage parasites (not shown) and are conserved in different genes, they may act as targets for negatively regulating repressor proteins, expressed at other stages of parasite development. Such a mechanism could result in silencing the expression of a variety of different genes while the parasite undergoes developmental changes.

## Materials and Methods

### Cultivation of Parasites

Plasmodium falciparum (strain A2) was grown and maintained as described (17,18). For the analysis of stage specific gene expression cultures were synchronized as described (8,12).

### Preparation of Nuclei

At a parasetimia of 10%, the contents of 50 10cm petri dishes were collected, harvested and washed once with 1x Trager's buffer. The following steps were carried out on ice. Erythrocytes were lysed in 0.05% saponin (19). The released parasites were sedimented at 10,000cpm for 10min. The parasite pellet was resuspended in 2ml of solution A (20mM PIPES pH 7.5, 15mM NaCl, 60mM KCl, 14mM beta Mercaptoethanol, 0.5mM EGTA, 4mM EDTA, 0.15mM spermine, 0.5mM spermidine, 0.125mM PMSF). 125ul of 10% NP-40 were added followed by vigorous vortexing for 10sec. The detergent was diluted out by the addition of 8ml of solution A. Nuclei were sedimented at 10,000rpm for 10min. Nuclei were transcribed immediately.

### In vitro Transcription of Nuclei

Nuclei were transcribed at 37°C for 10min in 600ul of solution B (50mM HEPES pH 7.9, 50mM NaCl, 10mM $MgCl_2$, 1.2mM DTT, 10mM creatine phosphate, 1mM GTP, 1mM CTP, 4mM ATP, 25% glycerol, 125units/ml rRNasin [Promega], 0.2mg/ml creatine kinase, 100ul

[$^{\alpha-32}$P]UTP 3000Ci/mmol. Radiolabled RNA was isolated as described below. Starting with 50 plates usually 20x10$^6$cpm were incorporated into RNA as determined by TCA precipition.

**Hybridization of radiolabelled RNA**

2pmol of plasmid DNA were denaturated at 95°C for 5min in a solution containing 0.1N NaOH and 0.5mM EDTA. 600ul of ice cold 10xSSC were added. The solution was filtered through nitrocellulose using a minifold device (Schleicher & Schuell). The air-dried filter was baked for 2hr at 80°C under vacuum. A prehybridization ( 50mM HEPES pH 7.4, 0.3M NaCl, 10mM EDTA, 0.2% SDS, 1mg/ml yeast tRNA, 0.5mg/ml poly(A), 1% sodium pyrophosphate, 5x Denhardt's solution without BSA) was carried out at 65°C over night. the prehybridization solution was replaced by 1.2ml of hybridization solution (50mM HEPES pH 7.4, 0.3M NaCl, 10mM EDTA, 0.2% SDS, 0.1% sodium pyrophosphate, 1x Denhardt's solution, 100ug/ml yeast tRNA, 100ug/ml poly(A) containing the radiolabelled RNA. After hybridization for 48hr at 65°C the filter was washed three times for 20min each in 2xSSC, 0.1%SDS at room temperature followed by two washed at 50°C in 0.1xSSC, 0.1% SDS. Filters were dried and exposed overnight at -70°C with screen.

**Isolation of RNA**

Total cellular RNA and radiolabelled RNA obtained upon <u>in vitro</u> transcription of nuclei were isolated by acidic guanidinium-

phenol chloroform method (2). Parasites were harvested from infected erythrocytes as outlined above. Volumes used to isolate radiolabelled RNA from nuclei are given in parenthesis. The parasite pellet was resuspended in 7ml (3ml) of solution D (4M guanidinium thiocyanate, 25mM sodium citrate, pH 7, 0.5% sarcosyl, 0.1M beta mercaptoethanol). Sequentially, 0.7 (0.3ml) of 2M sodium acetate pH 4, 7ml (3ml) of phenol, and 1.4ml (0.6ml) of chloroform-isoamyl alcohol mixture (49:1) were added. Upon centrifugation (10,000rpm 20min) RNA was precipitated from the supernatent by the addition of an equal volume of isopropanol. Radiolabelled RNA was resuspended in 1.2ml of hybridization solution. For the preparation of total RNA the pellet was resuspended in 0.5ml of TES buffer (20mM Tris pH 7.6, 2mM EDTA, 1% SDS, 200mM NaCl). 0.5ml of CsCl solution I (1g/ml CsCl, 25mM EDTA, 50mM Tris pH 8, 0.5% sarcosyl) was added and the mixture was carefully layed over a CsCl step gradient, composed of 2ml of CsCl solution II (5.7mM CsCl, 0.1mM EDTA pH8) overlayed by 1ml CsCl solution I. The RNA was sedimented at 35,000rpm for 16hr in a SW55 rotor. The final RNA pellet was resuspended in 200ul of TE buffer containing 0.1% SDS.

**Bacterial strains and libraries**

To minimize recombination and deletion events, plasmids carrying P. falciparum DNA were propagated in the multi recombinant deficient host, SURE (recB, recJ, sbcC201, uvrC, umuC::Tn5($kan^R$), mcrA, mcrB, mrr, lac, .(hsdRMS), endA1, gyrA96, thi, relA1,

supE44[F', proAB, lacI$^{q}$ZΔM15, Tn10(tet$^{r}$)]. SURE was purchased from Stratagene. All libraries used have been published previously (1,7). Libraries were screened using standard methods (7).

**Primer extension**

0.1pmol of endlabelled oligonucleotide primer (about 1.5x10$^{5}$cpm) and 30ug of RNA were coprecipitated. The precipitate was resuspended in 15ul of annealing buffer (40mM PIPES pH 6.4, 1mM EDTA, 0.4M NaCl, 50% formamide) and heated to 85°C for 10min. The annealing reaction was carried out at 20°C overnight. After ethanol precipitation the extension reaction was performed at 43°C for 1.5hr in a final volume of 20ul containing (50mM Tris/HCl pH 8.3, 75mM KCl, 3mM $MgCl_2$, 10mM DTT, 1mM each of the four dNTP, 40units of rRNasin [Promega], and 300 units of M-MLV H$^{-}$ reverse transcriptase [superscript, BRL]). Products were analyzed on a 6% denaturating acrylamide gel.

**S1 mapping**

Radiolabelled single stranded DNA as probes for S1 mapping were generated by extension of oligonucleotide primer annealed to denaturated plasmid DNA. Briefly, 10pmol of primer and 0.1pmol of plasmid DNA were mixed and denaturated in 0.2N NaOH, 0.2mM EDTA at 85°C for 5min. Upon ethanol precipitation annealing and extension were carried out as outlined in (11). The 5' end of

the DNA fragment was defined by digestion with a suitable restriction enzyme. The radiolabelled single stranded DNA fragment was purified by gel electophoresis. $2.0x10^5$cpm of the probe and 25ug of total cellular RNA were hyridized at 42°C overnight (in buffer described under primer extension). Digestion by S1 were carried out at 16°C for 1hr with 100units of enzyme (11) and the protected products were analyzed on a 6% denaturating acrylamide gel.

**Northern analysis**

5ug of total cellular RNA were fractionated on a 1% agarose-formaldehyde gel, transferred to nitrocellulose and hybridized with nicktranslated probes. For dot blot analysis of stage specific RNA accumulation, 10ug of total cellular RNA was absorbed to nitrocellulose using a minifold filtration device (Schleicher & Schuell) as decribed in (11). Hybridization conditions were performed as described (10).

**PCR amplification of genomic DNA**

Genomic DNA from the P. falciparum strain A2 was prepared as described (13). In a final volume of 100ul were mixed: 100ng of genomic DNA, PCR amplification buffer (10mM Tris/HCl ph 8.3, 50mM KCl, 1.5mM $MgCl_2$, 0.2mM each of the four dNTP), 25pmol each of the two primers , 2.5units of TAQ polymerase. PCR amplification was carried out in a automated termocycler (Perkins Elmer Cetus) for 35 cycles with the following settings: 1min 94°C, 2min 47°C,

3min 72°C. The amplified DNA was cut with internal restriction endonucleases and fractionated on a 1% low gelling agarose gel. Fragments of the correct size were cloned into pUC18 plasmids.

**Preparation of nuclear extracts**

Free parasites were prepared as described above. The method to prepare nuclear extracts was adapted from (16). Briefly, parasites were lysed and the nuclei were collected (lysisbuffer: 10mM HEPES pH 7.9, 10mM KCl, 0.1mM EDTA, 0.1mM EGTA, 1mM DTT, 0.5mM PMSF, 0.65% NP-40). Nuclei were lysed in 50ul of extraction buffer (20mM HEPES pH 7.9, 0.4M NaCl, 1mM EDTA, 1mM EGTA, 1mM DTT, 1mM PMSF). Per 10cm plate of culture and a parasetimia of 10% 20ug of nuclear proteins were obtained. 5ug of crude nuclear extract were incubated with 2fmol of double stranded, endlabelled oligonucleotides for 20min at room temperature. (bindingbuffer: 20mM HEPES ph 7.9, 100mM NaCl, 1mM EDTA, 1mM DTT, 5% glycerol, 0.25mg/ml BSA, ploy d(I/C) concentration as indicated; final volume: 15ul). The binding assays were analyzed by gel electrophoresis (4% acrylamide, 5% glycerol, 0.5xTBE).

## Acknowledgements

This work was supported by U.S. Army grant DAMD 17-89-Z-9003 and the World Health Organization. J.V.R. is a scholar in molecular parasitology supported by the Burroughs-Wellcome Fund.

Technical assistance was provided by Jahan Dadgar and Mark Samuels, and secretarial support was provided by Cynthia Ritter.

## Figure legends

**Figure 1. Transcriptional activity and RNA accumulation of blood stage genes.**

A. The transcription activity of the genes indicated was analyzed by nuclear run-on analysis. Nuclei were prepared from synchronized at cultures 18hr (ring), 30hr (trophozoite), and 41hr (schizont) after infection. Radiolabled RNA was analysed by hybridization to gene specific probes: rRNA (pPFrib1, (9), KAHRPII (14), GBP130 (7), HRPII (JVR, unpublished ), P195 (5), CS (p277-19), (4).

B. Total cellular RNA was isolated at the same time points as indicated above. 10ug of RNA each were analysed by northern analysis using the nicktranslated DNA probes (see above).

C. Prior to transcription nuclei prepared from a asynchronous growing culture were incubated with 100ug/ml of alpha amanitin for 15min on ice. Radiolabelled RNA was analysed as decribed above.

**Figure 2. Genomic organization and sequence of the GBP130 locus**

A. The GBP130 gene is linked at its 5' end to another transcriptional unit, called 3.8 gene. Both genes are transcribed in the same direction as indicated by arrows. Coding regions are shown as rectangular blocks. A flag and a stop sign indicate RNA initiation site of the GBP130 gene and the termination and polyadenylation signal of the 3.8 gene,

respectively. The genomic and cDNA clones obtained that span the entire locus are indicated.

**B.** Sequence of the intergenic region. The sequence is numbered with +1 referring to the RNA initiation site of the GBP130 gene, also emphasized by a flag. The transcribed regions are printed in bolded letters. A duplicaiton of 305bp 5' to the start point of transcription is boxed in. A sequence (position -114 to -141) which shares strong homologies to the SV40 enhancer is emphasized by a box with thick bars. G/C rich elements are indicated by smaller boxes. A box with a bold lower bar indicates the element TGTGTAC and a box with double lower bar the element TY(T/A)CCCT(T/G)T. Both elements are repeated five times. Their orientation is indicated by arrows.

**Figure 3. Characterization of transcriptional units within the GBP130 intergenic region.**

**A.** 5ug of total cellular RNA isolated from blood stage parasites was fractionated on a 1% agarose gel and analyzed by Northern blot hybridization. The position of the 28S and 16S RNA are indicated. A nicktranslated XbaI/NcoI fragment (X/N) taken from the 5' end of the intergenic region identifies a RNA species of 3.8 kb. The GBP130 cDNA clone, 2374, reveals the GBP130 transcript of 6.6kb. Upon cloning the XbaI/HindIII fragment of 2044 into pGEM3 strand-specific, radiolabelled RNA was generated (SP6, X/H). This probe identifies two RNA species of 6.6kb and 3.8bp confirming that the genes 3.8 and GBP130 are arranged in a

head-to-tail configuration.

B. Nuclear run-on analysis of the GBP130 intergenic region. Radiolabelled RNA derived from schizont nuclei was used as a probe for DNA fragments taken from the intergenic region. The DNA probes used are written above the autoradiogram and the radioactivity bound upon subtraction of the background below it. Two times less radioactivity hybridized to the probe H/H than to the proble identifying the 3.8 transcript and 8 times less, to the probe 8822 which contains the GBP130 coding region.

C. The RNA initiation site of the GBP130 gene was determined by primer extension. The primer used corresponds to a sequence from position +405 to +439. The length of the product was compared with a sequencing ladder of the genomic clone 8771 using the same primer. The first nucleotide transcribed is highlighted by a circle.

D. Schematic drawing of the GBP130 intergenic locus and the probes used for northern and nuclear run-on analysis.

**Figure 4. Upstream region of the KAHRP gene**

A. Schematic organization of the KAHRP upstream region. Relevant clones are indicated.

B. The sequence 5' of the KAHRP is shown. The sequence is numbered according to the start point of transcription, indicated by a flag. Interspersed G/C rich elements are indicated by boxes. An enhancer-like element is highlighted by a box with bold bars.

**Figure 5. Upstream region of the P195 gene.**

**A.** Schematic organization of the P195 upstream region, with relevant clones indicated.

**B.** Sequence of the P195 upstream region, +1 of the sequence is the RNA initiation site.

The initiation codon and the first amino acids are shown. An enhancer-like element is indicated by a box with thick bars. Boxes with broken bars reveal sequence elements that are conserved between the P195 gene of P. falciparum and its homolog Py230 of P. yoelii.

**Figure 6. Distribution of interspersed G/C rich and enhancer-like elements within putative plasmodial promoters**

Characteristic sequence elements of five plasmodial promoters are compared. GBP130, KAHRP and P195 are transcribed during the asexual blood stage of P. falciparum. The CS gene is expressed during the insect stage of P. knowlesi and the Py230 is the P. yoelii homologue of the P195. The putative promoter regions are aligned at the first transcribed nucleotide.

**Figure 7. Nuclear proteins form specific complexes with the conserved enchancer-like elements**

5ug of crude nuclear extract were incubated with 2fmol of oligo 1 or oligo 2 and analysed by gel electrophoresis. Oligo 1 contains a conserved enchancer-like element found in the putative promoter

region of the KAHRP gene and with some variations in the upstream region of the P195 gene. Oligo 2 contains a SV40 enhancer-like element which occurs in the promoter of the GBP130 gene. Both oligos form stable complexes with nuclear factors which can not be competed by polyd(I/C) (oligo 1: lanes 1 to 6; oligo 2: lanes 14 to 15). The polyd(I/C) concentration is indicated. Complexes formed between nuclear factors and oligo 1 were competed with 1 pmol of unlabled oligo 1 (lane 7, 1pmol of pUC plasmid DNA (lane 8), 1pmol of 10834 plasmid DNA (lane 9, KAHRP), 1pmol of 7364 plasmid DNA (lane 10, P195) and 1pmol of 8771 plasmid DNA (lane 11, GBP130). No complexes are formed between oligo 1 and 5 ug of protein extract, prepared from uninfected erythrocytes (lane 12). Lane 13 shows the free oligo 1. A random oligo does not bind any nuclear proteins (lane 16).

**Table 1. Promoter activity of blood stage genes.**
In vivo promoter activities are given in ppm by quantifying the nuclear run-on data. Radioactivity bound to the various DNA probes was determined by scintillation counting and divided by the total amount of radioactivity incorporated into RNA. The values given are the means of two independent experiments. The margin of error is 10%.

**Table 2. Comparison of conserved enhancer-like sequence elements**
**A. A sequence comparison between the SV40 enchancer and the sequences in the promoter region of the CS gene of P. knowlesi**

and of the GBP130 gene of P. falciparum is shown.

B. A comparison of elements found in the upstream region of the KAHRP and the P195 gene of P. falciparum and of the Py230 gene of P. yoelii. The genes P195 and Py230 contain two copies of the same element. A consensus sequence is shown.

## References


1. Certa, U., Ghersa, P., Dobeli, H., Matile, H., Kocher, H. P., Shrivastava, I. K., Shaw, A. R., and Perin, L. H. (1988). Aldolase activity of a Plasmodium falciparum protein with protective properties. Science 240, 1036-1038.

2. Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162, 156-159.

3. Corcoran, L.M., Thompson, J.K., Walliken, D. and Kemp, D.J. (1988). Homologous recombination within subtelomeric repeat sequences generates chromosome size polymorphisms in P. falciparum. Cell 53:807-813.

4. Enea, V. et al. (1984) DNA cloning of P. falciparum circumsporozoite gene: amino acid sequence of repetitive epitope. Science 225, 628-629.

5. Holder, A. A., Lockyer, M. J., Odink, K. G., Sandhu, J. S., Riveros-Moreno, V., Nicholls, S. C., Hillman, Y., Davey, L. S., Tizard, M. L. V., Schwarz, R. T., and Freeman, R. R. (1985). Primary structure of the precursor to the three major surface antigens of Plasmodium falciparum. Nature 317,

270-273.

6. Kimmel, A. R., and Firtel, R. A. (1982). The organization and expression of the Dictyostelium genome. In: The development of Dictyostelium discoideum (Loomis, W., ed.) pp. 233-244, Academic Press, New York.

7. Kochan, J., Perkins, M., and Ravetch, J.V. (1986). A tandemly repeated sequence determines the binding domain for an erythrocyte receptor binding protein of P. falciparum. Cell 44, 689-696.

8. Lambros, C., and Vanderberg, J. P. (1979). Synchronization of Plasmodium falciparum erythrocytic stages in culture. J. Parasitol. 65, 418-420.

9. Langsley, G., Hyde, J.H., Goman, M. and Scarfe, J.G. (1983) Cloning and characterization of the rRNA genes from the human malaria parasite P. falciparum. 11:8703-8717.

10. Lewis, P. A. (1990). Sequence analysis upstream of the gene encoding the precursor to the major merozoite surface antigens of Plasmodium voelii. Mol. Biochem. Parasitol. 39, 285-288.

11. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1989).

Molecular cloning. A laboratory manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

12. Pasvol, G., Wilson, R. J.M., Smalley, M. E., and Brown, J. (1978). Separation of viable schizont-infected red blood cells of Plasmodium falciparum from human blood. Ann. Trop. Med. Parasitol. 72, 87-88.

13. Pologe, L. G., and Ravetch, J. V. (1986). A chromosomal rearrangement in a P. falciparum histidine-rich protein gene is associated with the knobless phenotype. Nature 322, 474-477.

14. Pologe, L. G., Pavlovec, A., Shio, H., and Ravetch, J. V. (1987). Primary structure and subcellular localization of the knob-associated histidine-rich protein of Plasmodium falciparum. Proc. Natl. Acad. Sci. USA 84, 7139-7143.

15. Ruiz i Altaba, A., Ozaki, L. S., Gwadz, R. W., and Godson, G. N. (1987). Organization and expression of the Plasmodium knowlesi circumsporozoite antigen gene. Mol. Biochem. Parasitol. 23, 233-245.

16. Schreiber, E., Matthias, P., Muller, G. and Schaffner, W. (1989) Rapid detection of octamer binding proteins with minor extracts prepared from a small number of cells. Nuc.

Acids Res. 17:6419.

17. Trager, W. T., and Jansen, J. B. (1976). Human malaria parasites in continuous culture. Science 193, 673-675.

18. Trager, W. T., Tershakovec, M., Lyandvvert, L., Stanley, H., Lanners, N., and Grubert, E. (1981). Clones of the malaria parasite Plasmodium falciparum obtained by microscopic selection: their characterization with regard to knobs, chloroquine sensitivity, and formation of gametocytes. Proc. Natl. Acad. Sci. USA 78, 6527-6530.

19. Wallach, M. (1982). Efficient extraction and translation of Plasmodium falciparum messenger RNA. Mol. Biochem. Parasitol. 6, 335-342.

20. Weber, J.L. (1988) Molecular biology of malaria parasites. A review. Exp. Parasitol. 66:143-170.

21. Wesseling, J.G. et al. (1989) Stage specific expression and genomic organization of the actin genes of the malaria parasite P. falciparum. Mol. Biochem. Parasitol. 35:167-176.

**Table 1**

| GENE | PROMOTER ACTIVITY ppm | | |
|---|---|---|---|
| | R | T | S |
| rRNA | 120 | 125 | 110 |
| KAHRP | 20 | <1 | <1 |
| GBP130 | <1 | 16 | 17 |
| HRPII | 11 | 26 | <1 |
| P195 | 70 | 480 | 60 |
| CS | <1 | <1 | <1 |

Table 2

**A.**

```
SV 40       GCTGTGGAATGTGTGTCAG-TTAGGGTGTGGAAAGTCCCC
            * ****          * *** *** ******* **  ****
CS          GTTGTGAGTAAGCAG-CAGTTTAAGGTGTGGTAACCCCCC
                  ********  *    * ****** ****
GBP130      ATAAAATGTAAGCAGAAAAGGAATGGTGTGTTAACTTATT
```

**B.**

```
KAHRP       CATGTAGTGTAGTAATATTT
                 ***** * ***
P195(1)     ATATTAGTGCACTAAAGGAA
            * ** ********** * **
Py230(1)    AGATAAGTGCACTAA-G-AA
              ** ******** *  * *
Py230(2)    CGAAAAGTGCACCAT-G-AT
               ******** **
P195(2)     GATAAAGTGCATTAATTTTT
-------------------------------
Consensus   RA.AAGTGCACTAA.G
```

Fig. 1

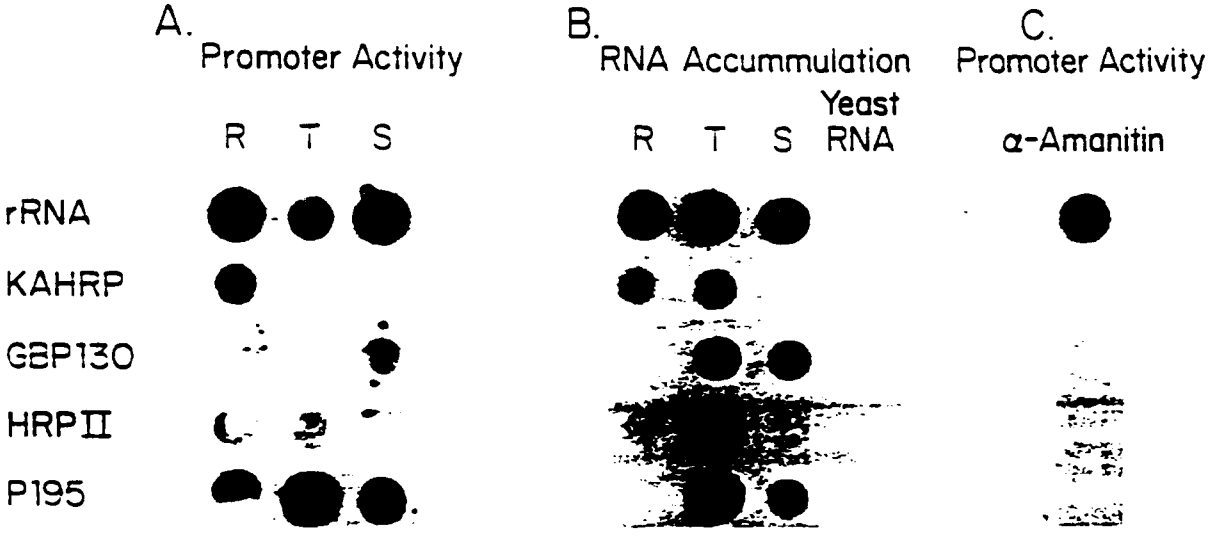

A.
Promoter Activity
R T S
rRNA
KAHRP
G2P130
HRPII
P195
B.
RNA Accummulation
R T S
Yeast RNA
C.
Promoter Activity
α-Amanitin

Fig. 2

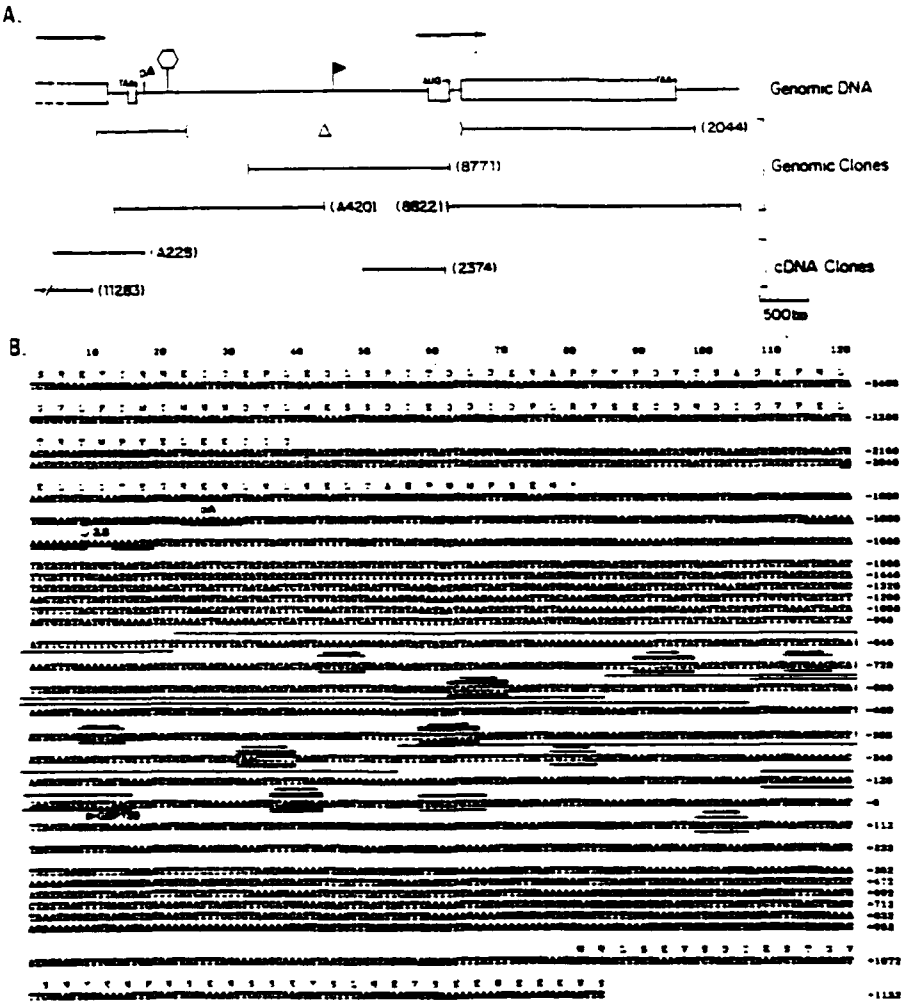

A.
Genomic DNA
(2044)
Genomic Clones
(8771)
(44201 (88221)
(1229)
(2374)
cDNA Clones
(11283)
500bp
B.

Fig. 3

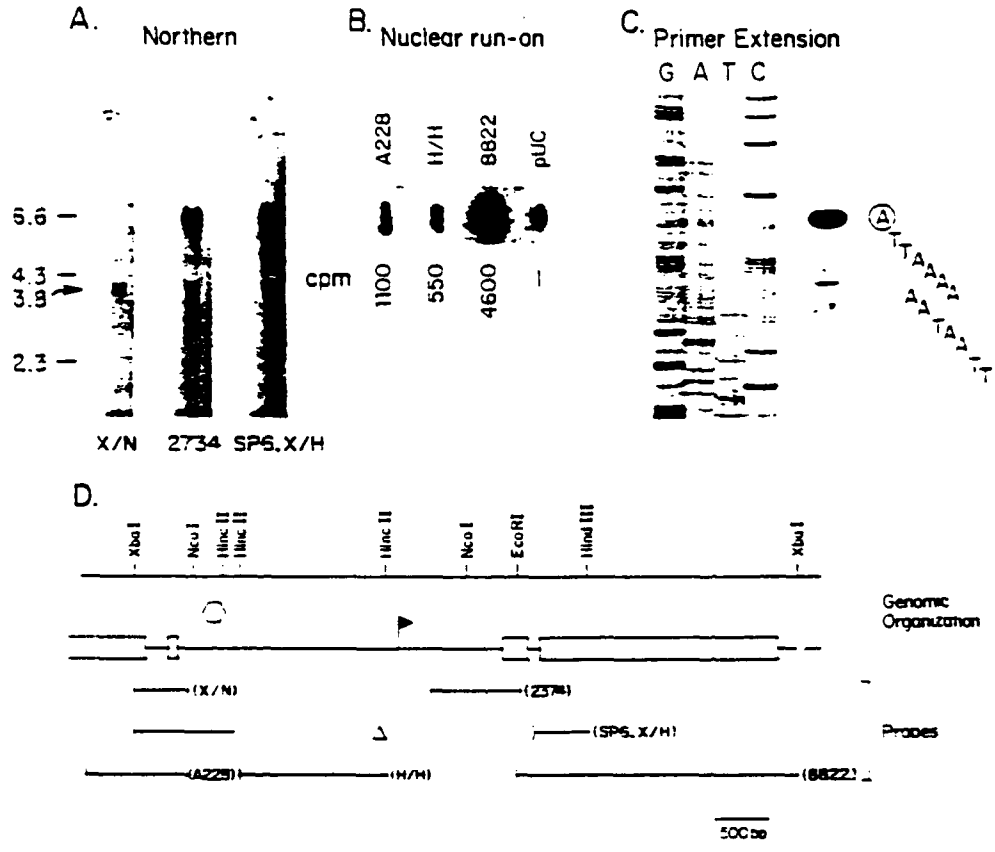
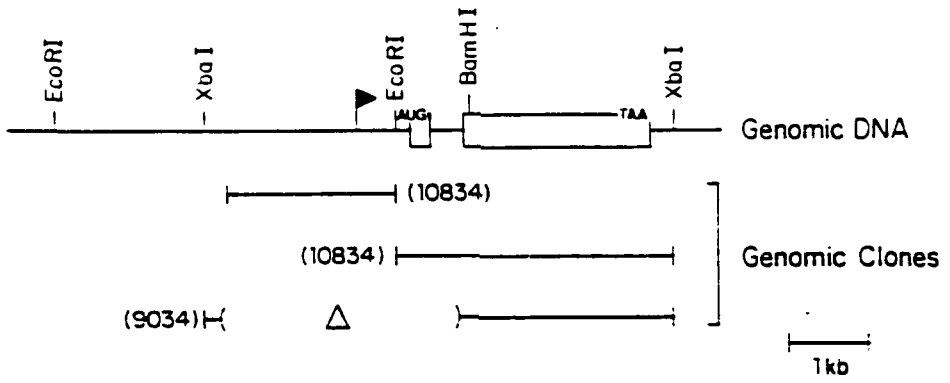

A.
Northern
B.
Nuclear run-on
C.
Primer Extension
G A T C
A228
H/H
8822
pUC
cpm
1100
550
4600
1
6.6
4.3
3.9
2.3
X/N
2734
SP6.X/H
D.
Xba I
Nco I
Hinc II
Hinc II
Hinc II
Nco I
EcoRI
Hind III
Xba I
Genomic Organization
(X/N)
(2734)
(SP6.X/H)
Probes
(A228)
(H/H)
(8822)
500 bp

Fig. 4

A.


EcoRI
Xba I
EcoRI
BamH I
Xba I
ATG
TAA
Genomic DNA
(10834)
(10834)
(9034)
Genomic Clones
1kb

B.

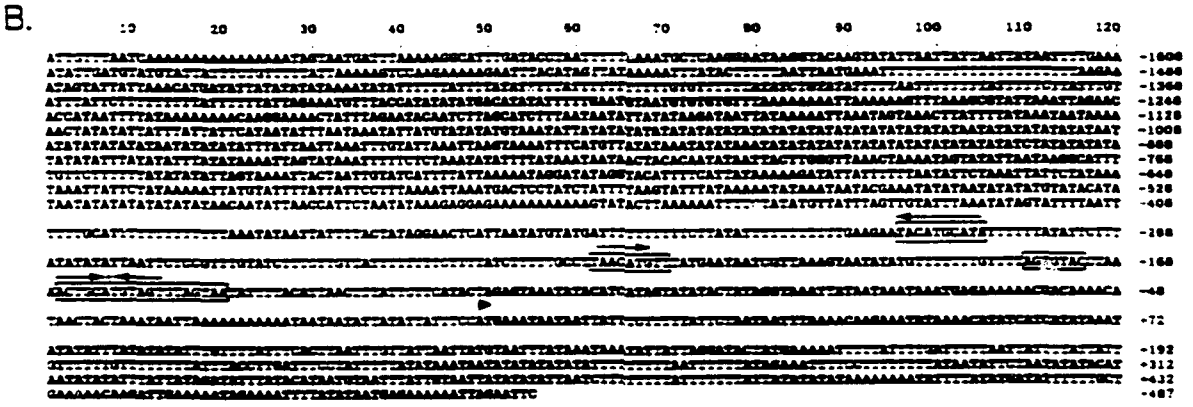


Fig. 5

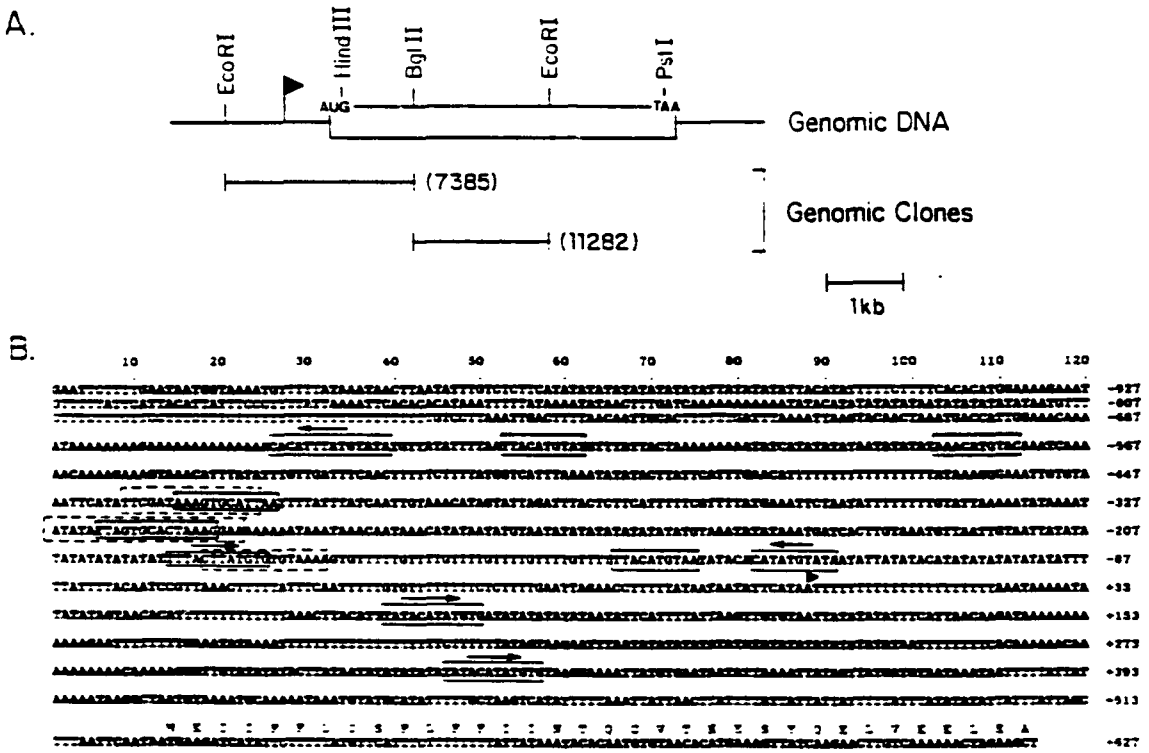
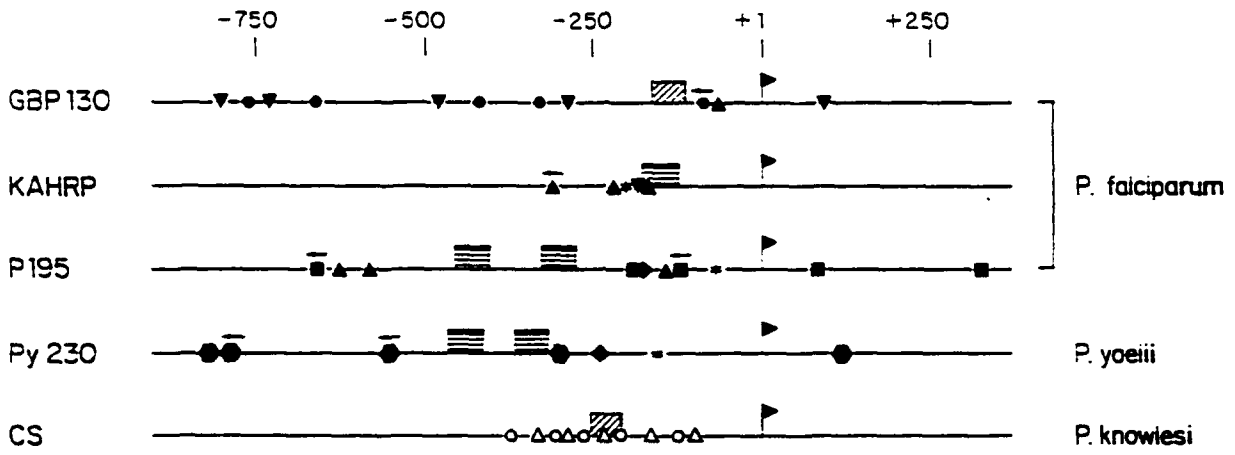
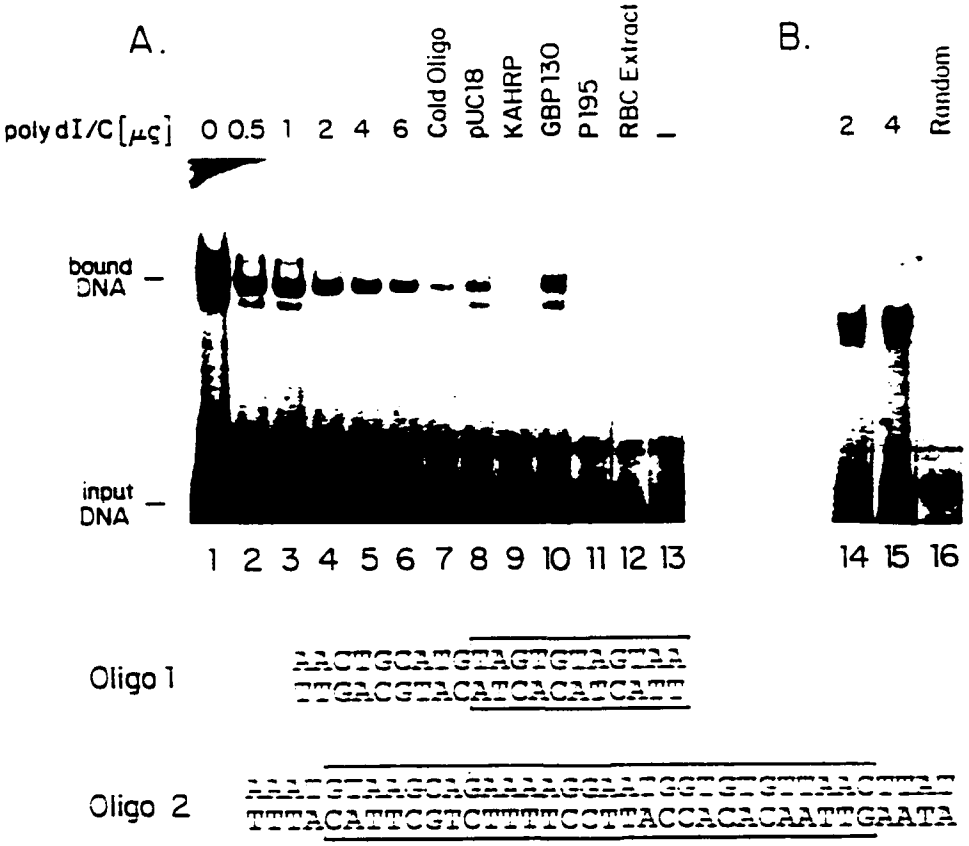

A.
EcoRI
Hind III
Bgl II
EcoRI
Pst I
ATG
TAA
Genomic DNA
(7385)
(11282)
Genomic Clones
1kb
B.

Fig. 6




SV40 enhancer-like
TY(T/A)CCCY(T/G)T
(T/A)GTGTAC
TRCATGTA
TATACATATGTG
TATATGCA•GTR
RA•AAGTGYAYTAA•R
TTRTGTGTGYAA
YCGTTAAA
TGCATGCA
CACCCCTC

Fig. 7


A.
B.
poly dI/C [μg] 0 0.5 1 2 4 6
Cold Oligo
pUC18
KAHRP
GBP 130
P 195
RBC Extract
–
2 4
Random
bound DNA –
input DNA –
1 2 3 4 5 6 7 8 9 10 11 12 13
14 15 16
Oligo 1
AACTGCATGTAGTGTAGTAA
TTGACGTACATCACATCATT
Oligo 2
AAATGTAAGCAGAAAAGGAATGGTGTGTTAACTTAT
TTTACATTCGTCTTTTCCTTACCACACAATTGAATA