Volume 16 Number 14 1988 Nucleic Acids Research

Mung bean nuclease exhibits a generalized gene-excision activity upon purified *Plasmodium falciparum* genomic DNA

Kenneth D.Vernick[1,2], Richard B.Imberski[2] and Thomas F.McCutchan[1]*

[1]Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892 and [2]Department of Zoology, University of Maryland, College Park, MD 20742, USA

Received March 30, 1988; Accepted June 8, 1988

ABSTRACT

A novel set of reaction conditions for mung bean nuclease has been described in which Plasmodium genes were specifically excised as intact fragments from purified DNA. We have now determined that under the new conditions mung bean nuclease cleaves precisely at sites outside of the coding region of every P. falciparum gene for which the extent of the protein coding region in genomic DNA is known. We conclude that this enzyme activity is probably a general one for P. falciparum genes. Introns are not specifically cleaved, although one gene contained a cleavage site within an intron. There is no direct relationship between dA•dT-richness and sites of cleavage under these conditions. Also contrary to the expectations of a model based on cleavage at denaturation bubbles, there was no general relationship between the concentration of the DNA denaturant, formamide, and the size of the resulting gene-containing fragments. Thus, the data strongly suggest the involvement of an altered DNA structure near gene boundaries in determining the recognition sites for this enzyme activity.

INTRODUCTION

Mung bean nuclease has typically been used for its single-stranded nuclease activity [1]. The sites of supercoil-dependent single-stranded nicking in prokaryotic DNAs have been characterized in an interesting series of studies [2-4]. However, it has been shown that under a different set of reaction conditions, an apparently new specificity of the enzyme precisely cleaves purified duplex DNA at sites which are outside of the coding region of genes from the malaria parasite Plasmodium falciparum [5]. The use of DNA similarly treated with mung bean nuclease has since proven of value in studies of trypanosomes [6], Leishmania [7], Toxoplasma [8], Giardia [9], and Plasmodium [10-15], although it was not known how the treatment worked. We now address several questions which were left unanswered at the time of the original report: 1) How general is this phenomenon? Do all or only a subset of the genes of an organism possess flanking mung bean nuclease recognition signals which will allow their precise excision under these reaction conditions. 2) Are there mung bean nuclease cleavage sites in introns? At the time of the original publication no introns had been described in DNA from malaria parasites. With the cloning and characterization of a number of Plasmodium genes containing introns, we are now able to address this question. 3) Does the mechanism of mung bean nuclease cleavage involve localized denaturation of dA•dT

rich regions of DNA and subsequent cleavage of the resulting single-stranded regions by the single-strand nuclease activity of the enzyme? Although initially attractive, this mechanism seemed unlikely because there was no direct relationship between dA•dT-richness and the location of cleavage sites.

To answer these questions, we have analyzed sites of mung bean nuclease cleavage in the vicinity of all of the 10 genes from P. falciparum for which the extent of the coding region in genomic DNA is known. These include genes expressed in different developmental stages of the parasite and genes both with and without introns. The enzyme activity which precisely excises genes under these conditions appears to be a general one for genes in the P. falciparum genome. For simplicity we now refer to this as the "genease activity" of mung bean nuclease. Work which is currently in progress suggests that there are significant differences in the enzymatic specificity of the nuclease under these conditions (in preparation). Introns are not recognized as discrete features by this activity, although it appears that some introns may contain recognition sites. The data do not support a model of cleavage at denaturated dA•dT rich regions, and point instead to the involvement of an altered DNA structure at distinct sites which border genes.

METHODS AND MATERIALS

Genomic P. falciparum DNA was prepared from cultured parasites of strain 3D7 [16] as described previously [17]. Mung bean nuclease reactions were performed essentially as previously described [5] but with several modifications. Routinely, 15 to 20 µg of genomic DNA were digested with 2 units of enzyme per µg DNA in a 100 µl reaction volume. Reaction components (excluding enzyme) were: 1) deionized water to 100 µl; 2) buffer of 0.2 M NaCl, 30 mM sodium acetate pH 4.6, and 1 mM $ZnSO_4$, kept as frozen 10X stock; 3) 1 mM dithiothreitol (BRL or Boehringer-Mannheim), diluted into reaction from 50 mM solution freshly made from frozen 1 M stock; 4) 20 µg/ml bovine serum albumen (Pentex fraction V), from frozen 1 mg/ml stock; 5) DNA; 6) concentrations of formamide (Clontech Ultrapure, deionized grade) of 30, 32.5, 35, or 40% vol/vol. These components were mixed and equilibrated at 50°C in a Lauda circulating water bath or equivalent for 5-10 min before the addition of enzyme.

Mung bean nuclease was routinely obtained from Promega Biotec or P-L Biochemicals. Some lots of enzyme provided as single-stranded nuclease did not exhibit genease activity (unpubl. observ.). Enzyme was diluted just before use to 10 units per µl in cold dilution buffer of 40% glycerol, 0.16 M NaCl, 30 mM sodium acetate pH 4.6, 1 mM $ZnSO_4$. Enzyme was added to the pre-equilibrated reaction at a concentration of 2 units per µg DNA. The reaction was mixed by pipetting and returned to the 50°C bath within 10 sec. After incubating for 15 min, the reaction was stopped with the addition of 3 volumes of cold 0.1 M Tris-HCl pH 8, 0.01 M EDTA. The reaction was extracted with phenol:chloroform (1:1) and

chloroform followed by ethanol precipitation of DNA in the presence of 2.5 M ammonium acetate.

Redissolved mung bean nuclease-treated DNA was electrophoresed in 0.8% agarose gels and transferred to nitrocellulose filters essentially as described [18]. For the experiments shown here, about 10 µg of digetsed DNA was loaded per lane to allow repeated probing and stripping of filters. Filters were probed with oligonucleotides made on an Applied Biosystems 380B synthesizer. About 0.2 pmol of kinased oligonucleotide was used per ml of hybridization solution, which was 6X SSC (SSC is 0.15 M NaCl, 0.015 M sodium citrate, pH 7), 0.2% SDS, 0.1% each bovine serum albumin, polyvinylpyrrolidone, and Ficoll, and 150 µg/ml sheared salmon sperm DNA. Hybridization was overnight at 42°C. Non-stringent washes were 3 x 15 min. at room temperature in 2x SSC, 0.1% SDS. One stringent wash was done for 3 min in 2x SSC, 0.1% SDS at a liquid temperature calculated by the function: T_w°C=[(2°C x dA•dT) + (4°C x dG•dC) - 5°C]. To strip probes from filters, boiling 0.1x SSC, 0.1% SDS was twice poured over filters which were rocked in a covered container for 15 min each time. Before re-probing a blot, the complete removal of the old probe was verified by taking an autoradiograph for the same length of exposure as the old probe.

For each gene, the formamide concentration which yielded the smallest fragment in a complete digestion was used. Completeness of the reaction was defined as resulting from the formamide concentration which produced a single predominant reaction product with the absence of submolar fragments smaller or larger than the major one. In general, the entire gene coding region is excised intact at formamide concentrations lower than those which generate complete digests, although the size of the fragment may be either larger or smaller than the fragment from a complete digestion, as discussed below. The exact conditions nescessary to give the smallest fragment containing an intact gene varies slightly between genes, but reaction conditions including 35% formamide seem to be of general use in P. falciparum. Depending on the particular gene, increasing the formamide above the effective concentration can result either in cleavage at less sensitive sites within the gene or else in nonspecific degradation of the sequence. In no case have we observed cleavages within a coding region at concentrations of formamide between 30 and 35% (unpubl. observ). Therefore, technical accuracy is of considerably more importance in reactions with higher percentages of formamide.

The sequences of the oligonucleotide probes were derived from published sequences as referenced for each gene in Results section, and were as follows, all 5'-3': RESA exon 1, 5' end, ATGAGACCTTTTCATGCATAT; exon 2, 3' end, GAACACAATGAAGAATATGAT. GBP-130, 5' end, ATGCGACTTTCTAAAGTATCT; 3' end, AATGCTGATAATAACGAAGCA. HSP-70, 5' end, GCTAGTGCAAAAGGTTCAAAA; 3' end, AATGCCAGGTGGTATGAATTTCC. Pfs-25, 5' end, TTTAAAAATGAATAAACTTTACAGTT; 3' end, TACATTATAAAAAAGCATACTGAAAA. CRA exon 1, 5' end, ATGAAAATCTTATCAGTATTT; exon 3, 3' end, CTCGTAAGTGGCCCTGAACAC. HRP-

1 exon 1, 5' end, GAAAGAAGGCTTTCCCTGTTT; exon 2, 5' end, ACTGCAATAATGGAAACGGAT; exon 2, 3' end, AGAAGAAAAGAGGATGCTGTG. DHFR-TS 5' end, ATGATGGAACAAGTCTGCGAC; 3' end, AATTCAATGGATATGGCTGCT.

RESULTS

Location of genease cleavage sites near P. falciparum genes

The extent of the protein coding region is known for a total of 10 genes from P. falciparum. Several strategies were used to determine whether the entire coding region of a gene was included in a mung bean nuclease-generated fragment. For 7 of the 10 genes, in which the complete nucleotide sequence of the coding region was known (RESA, GBP-130, HSP-70, Pfs-25, CRA, HRP-1, and DHFR-TS), a Southern blot of mung bean nuclease-digested genomic DNA was serially re-probed with oligonucleotides specific for the extreme 5' and 3' ends of the protein coding region. Since the same blot was hybridized with both probes, it could easily be determined whether the 5' and 3' ends of the gene resided on the same mung bean nuclease fragment. If probes from both ends of the protein coding region jointly hybridized to a mung bean fragment of adequate size to include the entire coding region, it was concluded that the intact gene was excised on a mung bean nuclease fragment. In several cases, additional probes were used which were at exon boundaries within the coding region.

For one gene (CSP), the sites of mung bean cleavage were previously determined by nucleotide sequence analysis in this laboratory. Finally, for two genes (gp195 and S-Antigen), we rely on published data generated from mung bean digestions carried out in other laboratories.

1) RESA

The Ring-infected Erythrocyte Surface Antigen (RESA) is specified by a relatively well-conserved gene expressed in blood-stage parasites [19]. The gene is composed of two exons, the first one of which encodes a 65 amino acid hydrophobic region which may be a signal sequence. The coding region in genomic DNA, including the intron, spans 3422 bp and has been completely sequenced. Probes specific for the 5'-end of exon 1 and the 3'-end of exon 2 both identified a common mung bean fragment 4.4 kb in size (Fig. 1A). Thus, both of the exons and the intervening intron were included in a single mung bean fragment.

2) GBP-130

The gene for the 130 kD Glycophorin Binding Protein (GBP-130) is expressed in merozoites and is conserved between parasite strains [20]. The entire protein coding region has been sequenced on cDNA clones and is 2322 bp. Most of the coding region has also been sequenced on clones of genomic DNA with the exception of the far 5' region. No introns were found in the sequenced portion, and no size discontinuities were observed between cDNA and genomic DNA restriction fragments in the far 5' region. Probes from the 5' and

3' ends of the protein coding region both identified a single mung bean fragment 4.0 kb in size (Fig. 1A). Since the GBP-130 gene contains repeated units of about 150 bp each which extend to the 3' end of the gene, a probe specific for the non-repeated 3' untranslated portion of the gene was made and was found to hybridize to the same band as the other probes (not shown).

3) HSP 70

The 70 kD heat shock protein (HSP 70) is expressed in heat-shocked blood stage parasites but not sporozoites and is encoded by a gene of 2043 bp, determined from a cDNA clone [21, 22]. The presence of introns is unknown since a complete genomic clone has not been sequenced, although introns are not a feature of the evolutionarily conserved class of proteins with which this HSP 70 shares extensive homology. Two probes were used, one specific for the 5' end and one for the 3' end of the protein coding region. Both identified a single mung bean fragment 2 kb in size (Fig. 1A). The smaller faint band identified by the 3' probe is almost certainly the closely related Glucose Regulated Protein, GRP [23], since a probe made to a region of GRP which diverges from the HSP-70 sequence identified mainly this band (not shown). The P. falciparum GRP gene has not yet been fully characterized [22].

4) Pfs 25

The gene for the 25 kD ookinete surface antigen (Pfs 25) is expressed throughout the sexual stages of the parasite in the mosquito host and is specified by a protein coding region of 651 bp [24]. Two probes were used, from the 5' and 3' ends of the coding region. Both identified a single 800 bp mung bean fragment (Fig 1A).

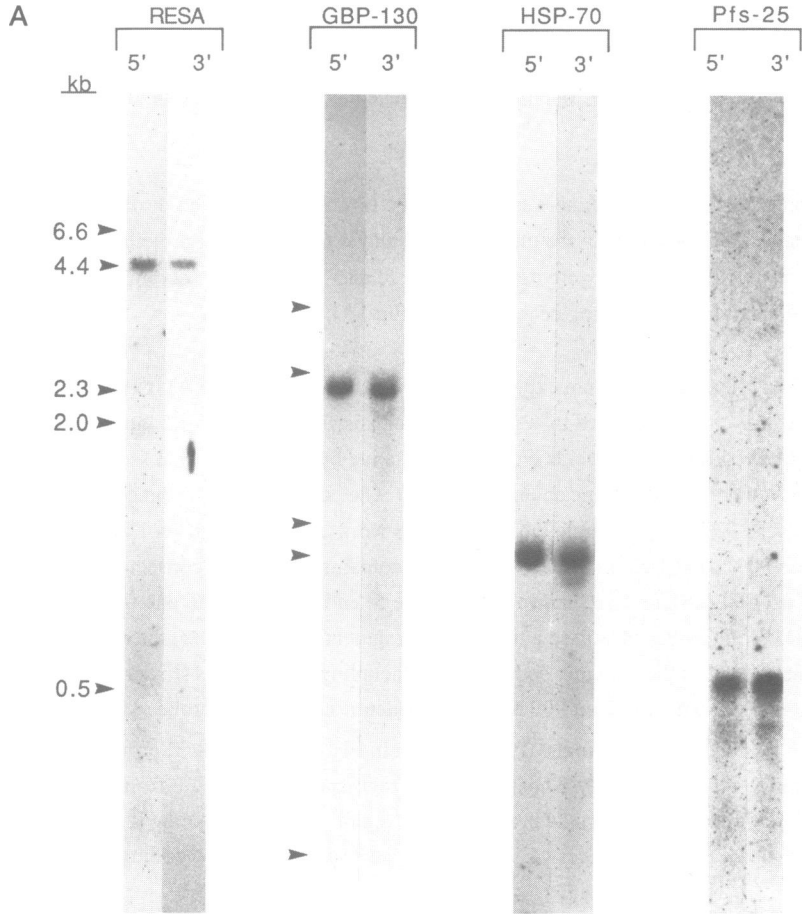
5) CRA

The 23 kD bloodstage protein exp-1 or Cross Reactive Antigen (CRA) is conserved between parasite strains and is specified by a gene which consists of 3 exons and 2 introns. The first exon encodes 40 amino acids including a putative hydrophobic signal sequence and exon 2 encodes a putative membrane anchor sequence. The coding region of the gene, including introns, spans 933 bp in genomic DNA [25]. It is worth noting that intron 1 is 85% and intron 2 is 90% dA•dT rich. Moreover, intron 2 contains a run of 56 consecutive dA•dT base pairs. Two probes were used, one specific for the 5' end of exon 1 and one for the 3' end of exon 3. Both probes identified a mung bean fragment 1060 bp in length, which is sufficient to contain the genomic coding region including both introns (Fig. 1B). A slightly smaller faint sub-band of about 900 bp was also identified by the 5' probe. On a different blot with less background than the 3' strip shown in Fig. 1B, the CRA 3' probe also hybridized to this minor band. Thus, the minor fragment does not result from a low frequency of cleavages within the protein coding region. We do not know if the 900 bp sub-band represents minor cleavages very close to either end of the 933 bp coding region, or

whether it represents a genomic sequence with strong homology to CRA.

6) HRP-1

The gene for the knob-associated histidine-rich protein (HRP-1, also called KP and KAHRP) is expressed in blood stage parasites and is encoded by 2 exons which, along with their intron, span 2328 bp in genomic DNA [26]. The first exon of 99 bp encodes a putative hydrophobic signal sequence. Three probes were used, one for the 5' end of exon 1, the 5' end of exon 2, and the 3' end of exon 2. The exon 1 probe identified a 450 bp mung bean fragment while probes from either end of the 1791 bp exon 2 hybridized to a single 2.2 kb fragment (Fig. 1B). Thus, all of exon 2 was contained within a mung bean fragment which was 400 bp longer than the exon. However, since the preceding intron is 438 bp long, there must be a cleavage within the intron which separated exon 1 from the main fragment. The intron is 95% dA•dT rich, including a run of almost 200 consecutive dA•dT base pairs.



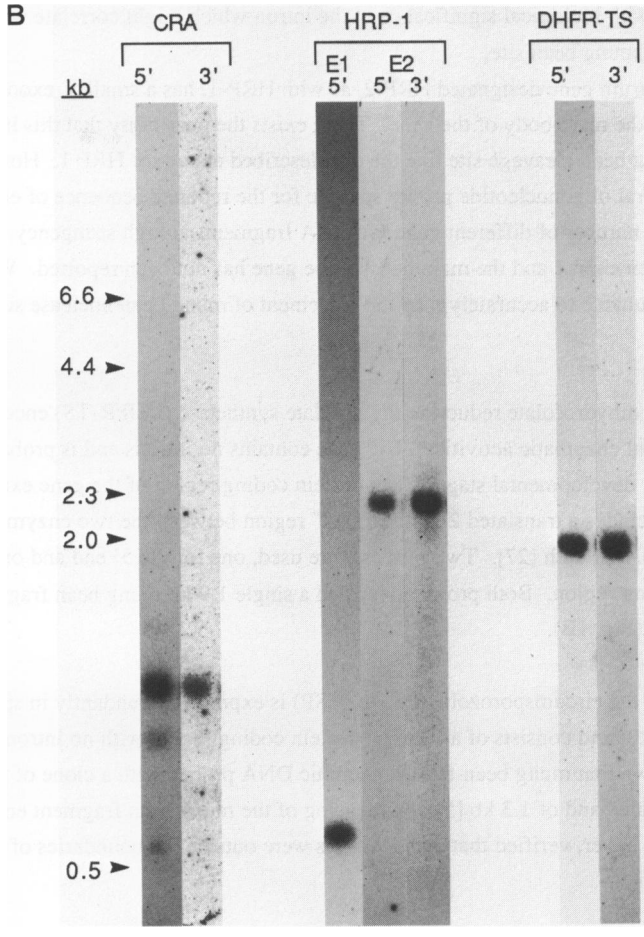


Figure 1. Sizes of gene-containing mung bean fragments. For each gene, a Southern blot of unrestricted, mung bean nuclease-treated genomic DNA from P. falciparum was serially hybridized with oligonucleotide probes specific for the far 5' and 3' ends of the protein coding region. Per cent formamide of the reaction is given in parentheses after each gene. **A)** A single blot was sequentially hybridized to both RESA (35%) probes. The positions of lambda/HindIII size markers are indicated. Probes for GBP-130 (35%), HSP-70 (35%), and Pfs-25 (30%) were all hybridized to a single blot. Arrows next to the GBP-130 strip indicate the positions of lambda/HindIII size markers for these three genes. **B)** A single blot was serially hybridized to probes for CRA (32.5%), HRP-1 (32.5%), and DHFR-TS (35%). For the HRP-1 gene, probes were for the 5' end of exon 1 (E1), and for the 5' and 3' ends of exon 2 (E2).

However, it was shown that a run of 100 consecutive dA•dT base pairs did not cut even with an additional 7.5% formamide above the concentration used in this experiment [5]. Since it is not known what signals regulate gene expression in Plasmodium, we do not know whether

there are sites with biological significance in the intron which might correlate with the location of this mung bean site.

The P.falciparum gene designated HRP-2, as with HRP-1, has a small 5' exon separated by an intron from the main body of the gene. There exists the possibility that this intron contains a mung bean cleavage site like the one described above for HRP 1. However, we found that several oligonucleotide probes specific for the reported sequence of exon 1 [12] hybridized to a number of different genomic DNA fragments at high stringency. Further, the distance between exon 1 and the main body of the gene has not been reported. We have therefore been unable to accurately map the placement of mung bean nuclease sites around this gene.

7) DHFR-TS

The gene for dihydrofolate reductase-thymidylate synthetase (DHFR-TS) encodes a protein with bifunctional enzymatic activities. The gene contains no introns and is probably expressed in all developmental stages. The protein coding region of the gene extends over 1824 bp and includes a translated 279 bp "linker" region between the two enzymatic moieties which is 81% dA•dT rich [27]. Two probes were used, one for the 5' end and one for the 3' end of the coding region. Both probes identified a single 1.9 kb mung bean fragment in genomic DNA (Fig. 1B).

8) CSP

The gene for the circumsporozoite protein (CSP) is expressed abundantly in sporozoites in the mosquito host and consists of a 1236 bp protein coding region with no introns. It was previously shown that mung bean-treated genomic DNA probed with a clone of the entire CSP displayed a single band of 1.3 kb [5]. Sequencing of the mung bean fragment ends, as reported in that paper, verified that the cleavages were outside the boundaries of the protein coding region.

9) S-Antigen

The size of the mung bean nuclease fragment homologous to a probe for the entire gene from the blood stage S-Antigen of the Wellcome strain was recently reported. The 1.9 kb mung bean fragment was found to compare favorably to the 1919 bp size of the protein coding region [15].

10) gp195

The size of the mung bean nuclease fragment containing the 195 kD merozoite surface glycoprotein (gp195) was previously reported [11]. In that work, about 3.3 kb of the gp195 protein coding region from the CAMP strain was sequenced from genomic clone a88, including the probable 5' start of translation. Assuming about 37 kD of protein encoded per kb of DNA, the entire coding region would be about 5.25 kb. In Wellcome strain DNA, the entire gp195 coding region was 4920 bp and contained no introns [28]. Although the gene displays considerable internal restriction polymorphism between strains, all strains tested share the

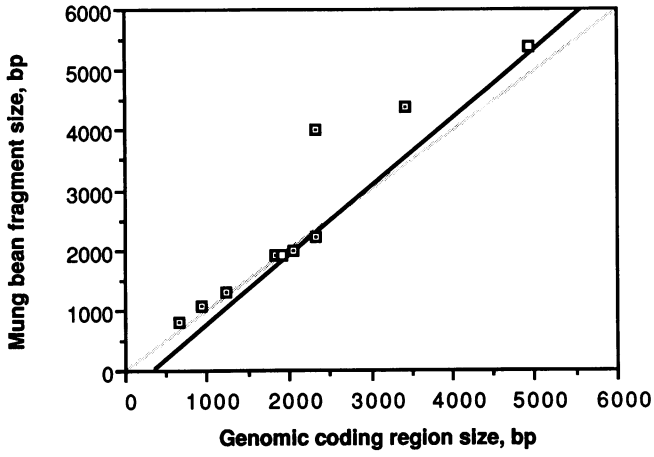


Figure 2. Size of mung bean fragment versus protein coding region. Data points from this laboratory are shown as dotted squares, data points from published accounts of other laboratories are shown as open squares. The solid black line represents the linear equation, y= -338+1.2x, derived from the data. The shaded line represents y=x, the expected result if mung bean fragment sizes were the same as the sizes of the coding regions.

same 4.9 kb EcoRI fragment which contains most of the gene [11]. Mung bean nuclease-treated genomic DNA from the CAMP strain, when probed with the a88 clone, gave a single band which was 5.4 kb. Since the upstream mung bean cleavage was near the translation start site, the resulting fragment would include enough sequence to code for a 195 kD protein.

Generality of genease specificity for P. falciparum genes

In all of the 10 genes from P. falciparum for which the extent of the protein coding region is known, mung bean nuclease cleaved P. falciparum DNA precisely and at positions outside of exons. In 9 of the 10 cases this resulted in single fragments containing intact genes, and the remaining case resulted in two fragments containing the two exons of the gene. The data from the preceding experiments is summarized in Fig. 2. The size of the mung bean fragment versus the size of the coding region was graphed for each gene, and the relationship between them was analyzed by linear regression and yielded the equation, y=-338 + 1.2x. The linear equation predicts that, at least in P. falciparum, the mung bean fragments which contain genes should on average be about 1.2 times the size of the protein coding region in genomic DNA.

Effect of formamide concentration

Interestingly, there does not appear to be a general relationship between the concentration of formamide in a mung bean reaction and the size of the resulting DNA fragments. This

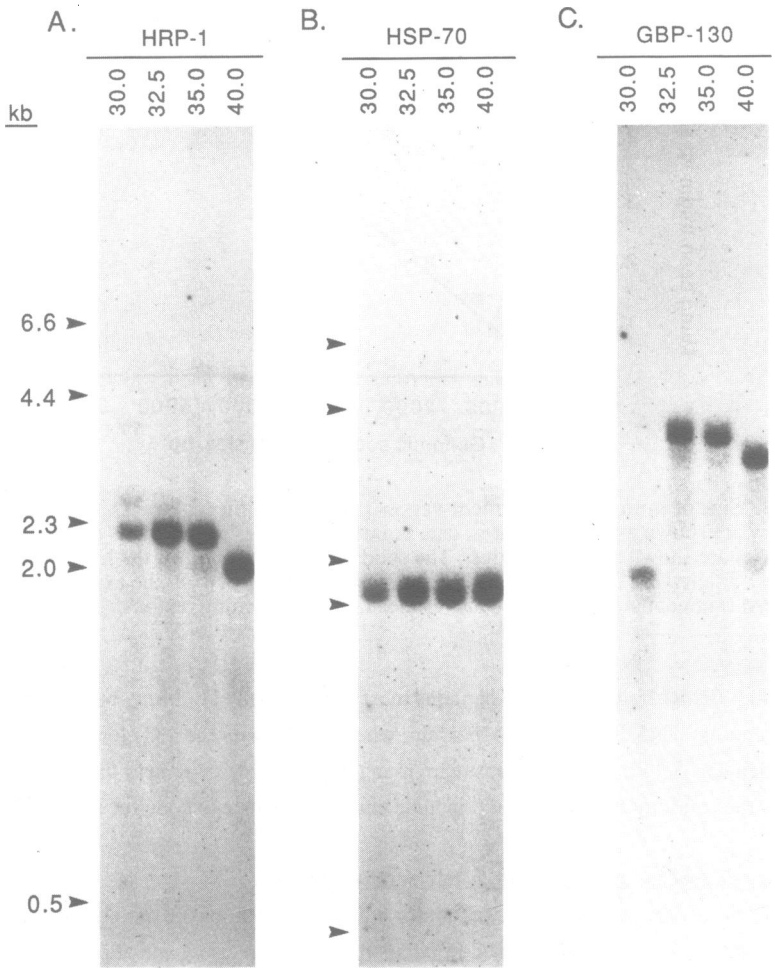


Figure 3. Effect of formamide concentration on mung bean fragment size. Samples from a single series of mung bean reactions on P. falciparum genomic DNA done in the presence of different concentrations of formamide were electrophoresed and blotted. Per cent formamide concentration in the reaction is indicated above each lane. Panels B and C were re-hybridizations of a single blot. Probes were: **A)** HRP-1, exon 2, 3' end **B)** HSP-70, 5' end **C)** GBP-130, 3' end.

point is illustrated by three representative examples in which DNA from a single series of mung bean nuclease reactions done with increasing concentrations of formamide was hybridized at high stringency to probes specific for three different genes. Figure 3A shows the series of reactions hybridized to the probe specific for the 3' end of HRP-1 exon 2. The probe for the 5' end of the exon gave the same result (not shown). For this gene, there

was an inverse relationship between formamide concentration and fragment size of the gene-containing bands. The 2.2 kb band was faint in the 30% reaction, probably because the reaction was incomplete and homologous sequences remained in higher molecular weight DNA. A faint precursor to the 2.2 kb band is visible at about 2.5 kb. In 32.5 and 35% formamide, the 2.2 kb fragment was the major product, followed by a discrete step to the 1.9 kb product in 40% formamide. In Figure 3B, the same range of formamide concentrations, from 30 to 40%, had no apparent effect on the size of mung bean fragments containing the HSP-70 gene. A 2.0 kb band was the only specific product detectable by the 5' probe in the identical reactions as above. The 3' probe for HSP-70 gave the same pattern and also gave a faint band from a closely related gene, as discussed above (not shown). Finally, Figure 3C shows the result of hybridizing the same blot as in Fig. 3B with a probe specific for the 3' end of the GBP-130 gene. Here, the product band was smaller in the reaction done with the least formamide and larger in reactions done with more formamide. Two other gene probes, Pfs-25 and DHFR-TS, also resulted in patterns where the gene-containing fragment seen at one formamide concentration was followed by a larger gene-containing fragment at a higher formamide concentration (not shown). In the case of GBP-130, the 30% formamide reaction yielded a band of about 2.2 kb which hybridized to probes from either end of the GBP-130 protein coding region as well as a probe just outside the 3' end of the coding region (5' coding and 3' untranslated probes not shown), which indicates that the small fragment contained the entire gene for GBP-130. At 32.5% and 35% formamide, the same three probes identified a 4.0 kb fragment on the same blot and at the same high hybridization stringency. Finally, at 40% formamide, the 3.8 kb product band hybridized to the two 3' probes but the 5' probe gave no signal, indicating that a cleavage had occurred inside the 5' end of the protein coding region. As mentioned earlier, increasing the formamide concentration above 35% can occasionally result in cleavage at less sensitive sites within the gene, although we have not observed such cleavages in any gene at concentrations between 30 and 35% (unpubl. observ). It is unlikely that differential cleavage of related genomic sequences is involved in generating the GBP-130 hybridization pattern shown here, since the gene was found by Southern blot analysis to be a single-copy gene conserved between strains and not related to other sequences in the genome [20].

DISCUSSION

We have found that, in all of the 10 genes from P. falciparum for which the extent of the protein coding region is known, mung bean nuclease cleaves purified, deproteinized genomic DNA at defined sites outside of protein coding regions. In 9 of the 10 cases, the major reaction product consisted of a single DNA fragment which contained the entire extent of the gene in genomic DNA, including any introns. In one case the gene-containing fragment was split by a specific cleavage within the intron of the gene into two fragments, one containing

each exon. The relationship between the size of the coding region and the size of the mung bean fragment was analyzed by linear regression. The resulting linear equation predicts that, at least in P. falciparum, mung bean fragments for genes should average about 1.2 times the size of the protein coding region in genomic DNA. We refer to "genease activity" to denote the enzyme specificity which yields predominantly gene-containing fragments. Studies of the enzymology of this apparently new specificity are currently in progress.

Three of the genes tested here, RESA, CRA, and HRP-1, include introns. The mung bean fragments obtained for RESA and CRA exhibited no specific cleavage in the one and two introns, respectively, which interrupt the coding regions of those genes. However, in the HRP-1 gene, a specific cleavage in the single intron isolated the 99 bp first exon and the 1791 bp second exon on independent mung bean fragments.

In the case of two other genes, HRP-2 [12] and HGPRT [13], DNA clones containing inserts which apparently terminated in an intron have been isolated in other laboratories from the original recombinant library of mung bean nuclease-treated P. falciparum genomic DNA [5]. These may well represent incidences of nuclease cleavage within an intron in a fashion similar to the HRP-1 gene. However, at the current time not enough information is available to make this determination. In neither of these two published cases were comparisons described between the cloned DNA and cleavage products in genomic DNA. Hence, the cloned inserts could be the result of, for example, minor nuclease reaction products or bacterial recombination. Further, since the original library was not methylated prior to the addition of EcoRI linkers, the termini of cloned inserts could also have resulted from cleavage at authentic EcoRI restriction sites.

It appears from these and other observations that introns themselves are not recognized by mung bean nuclease under these conditions as a distinct structural feature, although they can contain within them specific cleavage sites. The cleavages which do occur within certain introns may be informative by virtue of not being general events. In other systems such as the enhancer elements of immunoglobulin genes, sites of biological importance are known to reside within introns [29]. In Plasmodium, the cis-acting genetic signals which regulate expression remain uncharacterized.

The absence of any regular association between the concentration of formamide in a mung bean reaction and the size of the resulting DNA fragments has important implications for our understanding of the mechanism of gene excision by mung bean nuclease under these reaction conditions. One might imagine that the reaction proceeds by the denaturation of duplex DNA by formamide and temperature in dA•dT rich regions followed by enzymatic cleavage at unpaired bases in a putative denaturation bubble. If this model were true, it would necessarily follow that: 1) the location of mung bean cleavage sites should be directly dependent on local dA•dT richness, and 2) the size of mung bean fragments should generally be inversely related to the concentration of formamide in the reaction, since a higher

concentration of formamide should produce a more extensive region of denaturation.

It has been shown, though, that there is no direct dependence of mung bean cleavage under these conditions on dA•dT richness in DNA from Plasmodium [5], trypanosomes [6], or Leishmania [7], since in each of these cases sites which were cut were no more dA•dT-rich than nearby regions which did not cut. Moreover, a run of 100 consecutive dA•dT base pairs did not cut in less than 45% formamide [5] while the P. falciparum genes shown here were excised in reactions containing 30 to 35% formamide. In the present work, there were sizable runs of dA•dT base pairs internal to the mung bean fragments containing some of the P. falciparum genes. Two examples are: the 56 consecutive dA•dT base pairs in a 135 bp CRA gene intron which is 90% dA•dT rich [25]; and the 203 bp RESA gene intron which is 90% dA•dT and contains within it 5 runs of greater than 20 consecutive dA•dT base pairs each [19]. Since published primary sequences of sites of mung bean cleavage in Plasmodium [5, 10] are less dA•dT rich than this, there is no foundation in the data to suggest a direct relationship between dA•dT richness and mung bean cleavage sites.

We also found that the size of the gene-containing mung bean fragment does not necessarily decrease with increasing formamide in the reaction. In fact, in three cases (only one example shown in Fig. 3C) the size of the gene-containing fragments was seen to increase with an increase of formamide concentration. We are not aware of a way in which these observations regarding dA•dT richness and fragment size can be explained by a model of duplex melting in the conventional sense. It appears more likely that a discrete and locally-defined structural form in the DNA which flanks genes is revealed to the enzyme by the conditions in the reaction mixture. We do not yet know the structural characteristics of the sites recognized by mung bean nuclease under the genease reaction conditions used here. However, it has been clearly shown that there are altered DNA structures such as Z-DNA or poly dG•dC which can induce regions of unpaired bases in the sequence adjoining themselves in a way which is dependent on torsional stress but independent of the dA•dT content and the primary sequence of the adjoining DNA [30-31]. Thus, rather than promoting duplex melting, the genease reaction conditions may instead act to deform and torsionally stress linear DNA in a way which exposes structural forms at gene boundaries so they become recognizable by the nuclease. It seems unlikely that such well-defined sites generally positioned near gene boundaries would lack biological significance. Experiments are currently in progress which will help to describe this enzyme activity, the characteristics of genease recognition sites, and their biological meaning.

*To whom correspondence should be addressed

REFERENCES

1. Kowalski, D., Kroeker, W.D., and Laskowski, M., Sr. (1976) Biochemistry 15, 4457-4463.

2. Kowalski, D. (1984) Nucl. Acids Res. 12, 7071-7086.
3. Sheflin, L.G., and Kowalski, D. (1984) Nucl. Acids Res. 12, 7087-7104.
4. Sheflin, L.G., and Kowalski, D. (1985) Nucl. Acids Res. 13, 6137-6154.
5. McCutchan, T.F., Hansen, J.L., Dame, J.B., and Mullins, J.A. (1984) Science 225, 625-628.
6. Brown, K.H., Brentano, S.T., and Donelson, J.E. (1986) J. Biol. Chem. 261, 10352-10358.
7. Muhich, M.L., and Simpson, L. (1986) Nucl. Acids Res. 14, 5531-5556.
8. Johnson, A.M., Illana, S., Dubey, J.P., and Dame, J.B. (1987) Exp. Parasitol. 63, 272-278.
9. Adam, R.D., Aggarwal, A., Lal, A.A., de la Cruz, V.F., McCutchan, T.F., and Nash, T.E. (1988) J. Exp. Med. 167, 109-118.
10. Dame, J.B., Williams, J.L., McCutchan, T.F., Weber, J.L., Wirtz, R.A., Hockmeyer, W.T., Maloy, W.L., Haynes, J.D., Schneider, I., Roberts, D., Sanders, G.S., Reddy, E.P., Diggs, C.L., and Miller, L.H. (1984) Science 225, 593-599.
11. Weber, J.L., Leininger, W.M., and Lyon, J.A. (1986) Nucl. Acids Res. 14, 3311-3323.
12. Wellems, T.E. and Howard, R.J. (1986) Proc. Natl. Acad. Sci. USA 83, 6065-6069.
13. Sullivan, M.A., Lloyd, D.B., and Holland, L.E. (1987) In Agabian, N., Goodman, H., and Nogueira, N. (eds), Molecular Strategies of Parasitic Invasion, Alan R. Liss, New York, pp. 575-584.
14. McCutchan, T.F., de la Cruz, V.F., Lal, A.A., Gunderson, J.H., Elwood, H.J., and Sogin, M.L. (1988) Mol. Biochem. Parasitol. 28, 63-68.
15. Nicholls, S.C., Hillman, Y., Lockyer, M.J., Odink, K.G., and Holder, A.A. (1988) Mol. Biochem. Parasitol. 28, 11-20.
16. Walliker, D., Quakyi, I.A., Wellems, T.E., McCutchan, T.F., Szarfman, A., London, W.T., Corcoran, L.M., Burkot, T.R., and Carter, R. (1987) Science 236, 1661-1666.
17. Dame, J.B., and McCutchan, T.F. (1983) Mol. Biochem. Parasitol. 8, 263-279.
18. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
19. Favaloro, J.M., Coppel, R.L., Corcoran, L.M., Foote, S.J., Brown, G.V., Anders, R.F., and Kemp, D.J. (1986) Nucl. Acids Res. 14, 8265-8277.
20. Kochan, J., Perkins, M. and Ravetch, J.V. (1986) Cell 44, 689-696.
21. Bianco, A.E., Favaloro, J.M., Burkot, T.R., Culvenor, J.G., Crewther, P.E., Brown, G.V., Anders, R.F., Coppel, R.L., and Kemp, D. (1986) Proc. Natl. Acad. Sci. USA 83, 8713-8717.
22. Yang, Y., Tan-ariya, P., Sharma, Y.D., and Kilejian, A. (1987) Mol. Biochem. Parasitol. 26, 61-68.
23. Chappell, T.G., Welch, W.J., Schlossman, D.M., Palter, K.D., Schlesinger, M.J., and Rothman, J.E. (1986) Cell 45, 3-13.
24. Kaslow, D.C., Quakyi, I.A., Syin, C., Raum, M., Keister, D.B., Coligan, J., McCutchan, T.F., and Miller, L.H. Submitted.
25. Simmons, D., Woollett, G., Bergin-Cartwright, M., Kay, D., and Scaife, J. (1987) EMBO J. 6, 485-491.
26. Sharma, Y.D., and Kilejian, A. (1987) Mol. Biochem. Parasitol. 26, 11-16.
27. Bzik, D.J., Li, W., Horii, T., and Inselburg, J. (1987) Proc. Natl. Acad. Sci. USA 84, 8360-8364.
28. Holder, A.A., Lockyer, M.J., Odink, K.G., Sandu, 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) Nature 317, 270-273.
29. Bergman, Y., Rice, D., Grosschedl, R., and Baltimore, D. (1984) Proc. Natl. Acad. Sci. USA 81, 7041-7045.
30. Kohwi-Shigematsu, T., and Kohwi, Y. (1985) Cell 43, 199-206.
31. Kohwi-Shigematsu, T., Manes, T., and Kohwi, Y. (1987) Proc. Natl. Acad. Sci. USA 84, 2223-2227.