Volume 15 Number 8 1987 Nucleic Acids Research

Sequence and S1 nuclease mapping of the 5′ region of the dihydrofolate reductase–thymidylate synthase gene of *Leishmania major*

Geoffrey M.Kapler[1,2], Kang Zhang[1] and Stephen M.Beverley[1]*

[1]Department of Pharmacology and [2]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA

Received October 28, 1986; Revised and Accepted March 12, 1987 Accession no. Y00124

ABSTRACT

The 5' structure of mRNA transcribed from the dihydrofolate reductase-thymidylate synthase (DHFR-TS) gene of the protozoan parasite *Leishmania major* has been characterized. S1 nuclease mapping identifies a heterogenous 5' structure which is not affected by growth phase or developmental stage. The DNA sequence of the 5' region of the DHFR-TS gene does not reveal homology with other trypanosomatid genes, eukaryotic consensus genetic elements, or the mammalian DHFR promoter element. This latter finding is especially significant as we show that the 5' region of the *E. coli* DHFR gene exhibits homology to the mammalian DHFR promoter element, despite their greater evolutionary distance.

INTRODUCTION

Protozoan parasites of the family Trypanosomatidae exhibit a number of novel molecular phenomena (1,2). Among these is the unusual structure of cellular mRNAs, which exhibit a bipartite structure in which a common 35 nucleotide extension is found on the 5' end (3,4). This sequence, the "mini-exon" (5) or "spliced leader" (6), is encoded by a separate genetic locus from the body of the mRNAs (7,8), and is added by a mechanism(s) which is unknown. Another feature of trypanosomatid genes has been the general absence of "consensus" genetic elements implicated in various aspects of gene expression in other eukaryotes (reviewed in 1). To approach some of the questions associated with gene expression in trypanosomatids, we reasoned that examination of ubiquitous genes which have been studied in other species would provide useful comparative insights.

The enzymes dihydrofolate reductase (DHFR) and thymidylate synthase (TS) have been extensively studied in many species, because they catalyze critical steps in the de novo synthesis of precursors for DNA synthesis and are important targets for

© IRL Press Limited, Oxford, England.

chemotherapy (9,10). In mammalian cells the expression of the genes encoding these enzymes is growth phase and cell cycle regulated at both transcriptional and post-transcriptional steps (11-17). These genes have been examined using DNA transfection methodology, allowing localization of cis-acting transcriptional elements such as promoters, etc. (18,19). Structural studies indicate heterogeneity at both the 5' and 3' ends of DHFR and TS mRNAs (15,18,20,21,22), with the profile of the different mouse DHFR mRNAs being invariant thoughout the cell cycle (23). DHFR and TS thus represent paradigms for the study of gene structure and expression.

In *Leishmania* and all protists DHFR and TS exist as a bifunctional polypeptide encoded by a single gene (24-26). We and others have recently reported the DNA sequence of the protein coding portion of the DHFR-TS gene (27,28), which consists of a 5' DHFR domain followed by the TS domain. These studies have been aided by the availability of lines which have amplified the DHFR-TS gene on a 30 kb extra-chromosomal circular DNA (29) and consequently overproduce the DHFR-TS mRNA and protein. In this paper we report on the 5' structure of the *Leishmania* DHFR-TS gene and mRNA.

MATERIALS AND METHODS

Preparation of promastigotes and amastigotes.

The wild-type LT252 line of *Leishmania major* (30,31) and its methotrexate (MTX) resistant derivative R1000-3 line (29) were studied. Promastigotes were propagated in vitro in medium M199 as described (25), which contained 1 mM MTX for the R1000 line. Stationary phase cells were harvested 48 hours after reaching maximal cell density. The viability of these cells was >95% as determined by light microscopic examination of morphology and motility.

Intracellular amastigotes were obtained from lesions initiated by intradermal injection of 2×10^7 stationary phase promastigotes into the dorsal hind foot of Balb/c mice. Amastigotes were purified from the infected tissue by Percoll gradient centrifugation (32).

RNA isolation.

Promastigote and amastigote RNA was isolated by lysing parasites in 5 M guanidinium isothiocyanate followed by precipitation of RNA with 4 M lithium chloride (33). Polyadenylated RNA was prepared by oligodeoxythymidylate cellulose chromatography (34).

RNA filter hybridization.

RNA was fractionated on 6% formaldehyde, 20 mM MOPS (pH 7.0), 1% agarose gels (35). Following electrophoresis the gels were treated with 50 mM NaOH, 10 mM NaCl for 45 minutes; 100 mM Tris-HCl (pH 7.4) for 45 minutes; and 20x SSPE for 1 hour (1x SSPE= 10 mM $NaPO_4$ [pH 7.0], 0.15 M NaCl) and blotted overnight onto nitrocellulose filters. Filters were subsequently baked under vacuum for 3 hours at 80°C and stained with methylene blue prior to hybridization (35), allowing quantitation of RNA and identification of molecular weight markers. Filters were prehybridized for 12 hours at 45°C in 50% formamide, 0.75 M NaCl, 0.15 M Tris (pH 7.0), 0.1% SDS, 0.1% sodium pyrophosphate, 0.2% bovine serum albumin, 0.2% polyvinylpyrollidone, 0.2% ficoll, 10 ug/ml poly-U, 10 ug/ml poly-C, 100 ug/ml denatured salmon sperm DNA, followed by hybridization for 40 hours in the above solution containing 10^7 cpm/ml of denatured radiolabeled probe. The hybridization probe consisted of a 0.9 kilobase (kb) Eco RI- Bgl II fragment, corresponding to the genomic sequence from -252 to +668 of the DHFR-TS gene (Fig. 1D). This fragment was radioactively labeled by randomly primed synthesis (36) to approximately 5 x 10^9 cpm/ug. Filters were washed for 10 min with 50% formamide, 5x SSPE, 0.1% SDS at 45°C, followed by four 20 minute washes at 55°C with 0.1x SSPE, 0.1% SDS, dried and exposed to X-ray film.

S1 nuclease mapping.

S1 nuclease analysis was performed according to the method of Berk and Sharp (37). Two probes were utilized: 1) the 575 base Sau 3A - Pst I restriction fragment (positions +135 to -440 relative to the AUG translation initiation site; Fig. 1D), labeled at the 5' Sau 3A site using polynucleotide kinase (35) and gamma ^{32}P-ATP (ICN); 2) the 458 bp Pvu II fragment (positions

+13 to -445), uniformly labeled on the mRNA-complementary strand by primed synthesis (M13 17 bp sequencing primer; New England Biolabs) in the presence of ^{32}P-dCTP (Amersham), using single stranded DNA from an M13 recombinant containing this fragment in the appropriate orientation (38). Polyadenylated RNA was annealed with the appropriate DNA probe for 3 hours at 53°C, digested with S1 nuclease and analyzed by electrophoresis on 4% acrylamide/8M urea gels (39). The fragment sizes reported are the averages of at least four determinations, rounded to the nearest 10 nucleotides.

DNA sequencing.

DNA sequencing was performed using the dideoxynucleotide technique (40) with alpha ^{35}S-dATP (41) and *Leishmania*-derived DNAs cloned into M13 vectors (38). The sequence from -760 to +1 was determined on both strands, while that from -952 to -761 was obtained solely in the mRNA-coding direction.

RESULTS

Blot analysis of the DHFR-TS mRNA.

Total RNA from the R1000 line reveals a single RNA species of 3.2 kb when hybridized with the 0.9 kb Eco RI- Bgl II fragment from the 5' end of the *Leishmania major* DHFR-TS gene (Fig. 1A, lane 1; see Fig. 1D for the location of the probe). This 3.2 kb transcript is present in polyadenylated RNA (Fig. 1A, lane 2) and absent in nonpolyadenylated RNA (Fig. 1A, lane 3). Additional hybridizing material evident in the total RNA, but not in polyadenylated or nonpolydenylated RNA is attributed to DNA contaminating the preparation; its apparent intensity reflects the relative abundance of the amplified DHFR-TS gene within the small *Leishmania* genome (25). The 3.2 kb polyadenylated RNA is the only RNA recognized by hybridization probes from throughout the DHFR-TS gene in log or stationary phase promastigotes or amastigotes (27; data not shown).

S1 nuclease mapping of DHFR-TS mRNA from promastigotes and amastigotes.

S1 nuclease analysis using the 575 bp Sau 3A- Pst I fragment (see Fig. 1D for location of probes) and polyadenylated RNA from the R1000 line revealed five protected fragments of approximate

sizes 330, 380, 430, 510 and 570 nucleotides (nt; Fig. 1B). The two major S1-resistant products map to positions -375 and -245 relative to the start of translation. Two minor products map to positions -295 and -195. In addition, a product the size of the fully protected Sau 3a- Pst I fragment was observed which was not detected when RNA was omitted (Fig. 1B, lane 1). S1 mapping of log and stationary phase R1000 promastigote mRNAs generated identical nuclease protection profiles (Fig. 1B , lanes 2 and 3) with the -245 product being approximately two-fold more abundant than the -375 product.

We next examined mRNAs isolated from the wild-type LT252 line, from either promastigotes propagated in vitro, or amastigotes isolated from chronic mouse lesions. Figure 1C shows the results of an S1 nuclease analysis using a uniformly labeled probe complementary to the DHFR-TS sense strand from +13 to -445. Protected fragments of approximate sizes 390, 310 and 260 nt are evident in the R1000 line (Fig. 1C, lane 1) and similarly sized fragments are evident in the promastigotes from the parental wild-type LT252 line (Fig. 1C, lane 2). Fragments corresponding to the -195 S1 site and fully protected probe were not detected with this probe. Only the 390 nt and 260 nt fragments were detected in the LT252 amastigote preparation (Fig. 1C, lane 3), however, the signal intensity would be insufficient to reveal less abundant species. These data indicate that by S1 analysis the structure of the 5' end of DHFR-TS mRNAs is the same in amplified and wild type promastigotes and wild type amastigotes. In addition, the levels of the S1 products are all elevated in the amplified R1000 line.

DNA sequence of the 5' region of the DHFR-TS gene.

Blot hybridization analysis of RNA revealed that the 0.7 kb Sal I- Eco RI fragment immediately 5' of the Eco RI site (Fig. 1D) hybridized to both the DHFR-TS mRNA as well as to another similarly abundant polyadenylated RNA (unpublished data). This suggested that the Sal I - Eco RI fragment must contain any non-transcribed DNA separating the two mRNAs, which might encode cis-acting transcriptional elements. We determined the nucleotide sequence of this fragment, extending 952 bp 5' of the DHFR-TS translation initiation codon (Fig. 2).

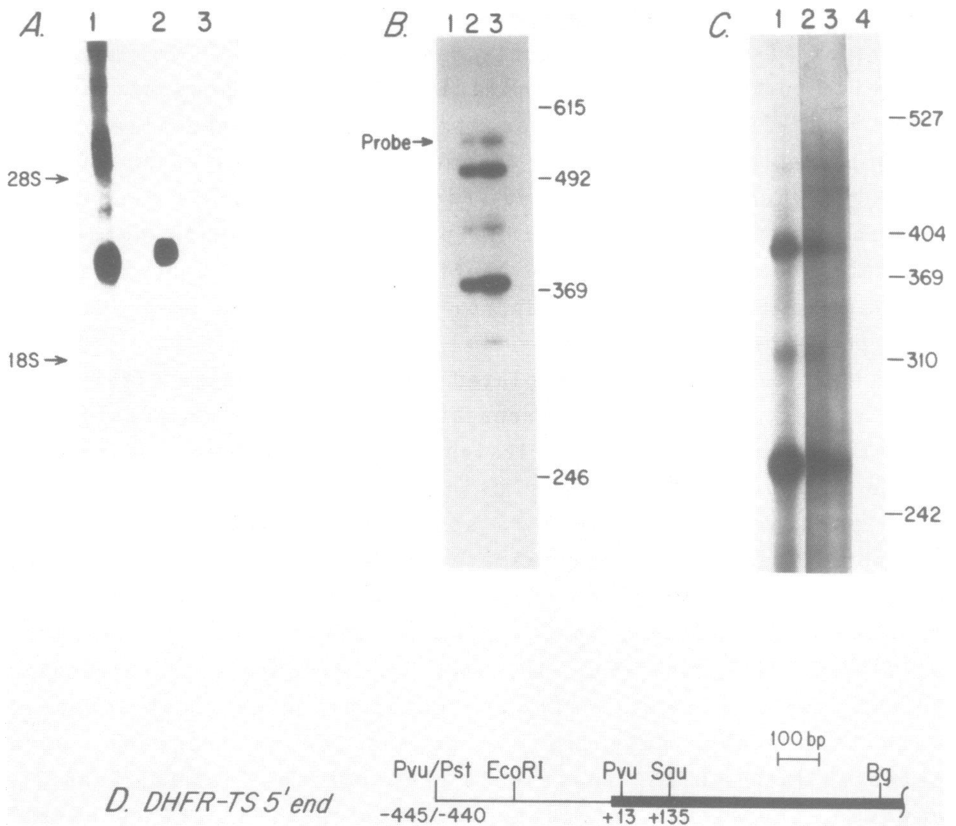

Figure 1. (A) Northern blot analysis of DHFR-TS mRNA. R1000 promastigote RNA was electrophoresed in a 1% agarose, 6% formaldehyde gel, transferred to nitrocellulose and hybridized to a radiolabeled 0.9 kb Eco RI - Bgl II fragment specific for the DHFR-TS gene (Fig 1D). Lane 1: 10 ug total RNA. Lane 2: 200 ng polyadenylated RNA. Lane 3: 8 ug nonpolyadenylated RNA. The positions of mouse 28S (5 kb) and 18S (2 kb) ribosomal RNAs are indicated by arrows. Hybridizing material in lane 1 greater than 3.2 kb in size is attributed to DNA contamination. (B) S1 nuclease mapping of the 5' end of promastigote DHFR-TS mRNAs. The 575 base Sau 3a- Pst I fragment 5' end labeled at the Sau 3A site, was hybridized to 1 ug of polyadenylated RNA isolated from log or stationary phase R1000 promastigotes and digested with S1 nuclease as described in Materials and Methods. Lane 1: no RNA. Lane 2: R1000 log phase RNA. Lane 3: R1000 stationary phase RNA. The mobility of DNA molecular weight markers (nt) are shown. (C) S1 nuclease analysis of amastigote form DHFR-TS mRNAs. Single-stranded uniformly labeled DNA corresponding to the 458 bp Pvu II fragment (positions +13 to -445) complementary to the DHFR-TS mRNA was used. Lane 1: 1 ug polyadenylated RNA from log phase R1000 promastigotes; lane 2, 1 ug polyadenylated RNA from wild-

type LT252 log phase promastigotes; lane 3, 0.3 ug polyadenylated RNA from wild-type LT252 amastigotes; lane 4, no RNA. The position of DNA molecular weight markers are shown. Lanes 2 and 3 were exposed to X-ray film four times longer than lane 1. (D) Restriction map of the 5' region of the *Leishmania major* DHFR-TS gene. The locations of relevant restriction enzyme cleavage sites are indicated. The protein coding region is represented by a thick black bar. The Sal I site referred to in the text is located at position -952.

We first consider the DNA sequences surrounding the DHFR-TS 5' termini revealed by the S1 analysis. The dinucleotide AG is found adjacent to the site of addition of the mini-exon to mRNAs transcribed from other kinetoplastid genes (42-45) and in higher eukaryotes marks the 3' border of intron sequences removed during RNA splicing (46). AG dinucleotides are in close proximity to the -375, -245 and -195 S1 sites, and preliminary data from primer extension analysis (not shown) suggests the addition of mini-exon sequences to one or more of the mRNAs mapping to the sites identified by S1 mapping. The sequence upstream of these sites do not otherwise exhibit significant sequence homology to one another. A sequence related to the eukaryotic splice branch acceptor sequence (YYYRAY; 47) is found at position -266. Interestingly, homology with the *L. major* mini-exon (sequence provided by J. Miller) was observed at two positions immediately 5' of the -375 site (11/13 and 10/12 nt), starting at positions -402 and -392 (Fig. 2). Homology with the mini-exon RNA has also been reported for other kinetoplastid genes (43).

The regions 3' of the major -375 and -245 S1 sites exhibit a limited degree of sequence homology, containing a pyrimidine-rich region with the conserved hexanucleotide sequence CTCCTC 20 and 18 bp 3' of the respective S1 endpoints. This region also exhibits comparable homology to the corresponding region of the beta-tubulin gene of *L. major* (J. Miller, unpublished data) and alpha- and beta-tubulins of *L. enrietti* (45). For all sequences, 10-20 bp downstream of the presumed or actual mini-exon addition site a conserved CTCYTC sequence is found, which is preceded by a pyrimidine-rich region. The significance of this observation is uncertain, as we can find this sequence at eight other positions throughout the DHFR-TS mRNA, including twice within the protein-coding region.

```
        SAL I
        GTCGACGGGGTGATGGAGAGAATGCTTGACCATGAGGACGCCGATGGGCGATGCGCCAGCGCGTGCTTGCGT

-880
ACTTCAATCGTTTTTCATCTTGATTCAGAGCTAAGTAGTAGCACACCACAGCAAGCGTAATCATGTAGCATGTATGCAGG

-800
TGCACGAGTAGGCACTCTCTGTCTCTGTCCGGCATCTATGATGCTCTCCGGCCGCGTGCGCGCGACGCCTCTGTGTGAGT

-720
GCTCGTGTGAGTGTGTGAGAGCTGTCCAGGCCTTTCCTCTGCCTACTCGTCCCACGCTCAGGCAGGCACAGAGGAATGAA

-640
AAGGAAGACGAGGACGCAACGCCAGGCGCATCGACATCGCGCACGTGTTTGCGGTGTACTGTGCGTCGATGATCGAGCGG

-560
CCAAGCGTCCAGCTTCACGCGAGGAGGACTGCATCCCTGCTCCTGTGCGCGTGCGTGCACGCGTCAAGGCACGCTCGATT

-480
CGTTCCCAGGGTGGATACCCTCCCCCCTCCCCCAGCTGCAGACAACCCCCACCCTCTCGAGGCTTGCACCCAGGCTCGTC

                 b
-400  a
AGTCTCTCTACCACTTTCTGCCTTCTGTAGTGCACCCCACCCCTGCATTCTCCTCTCTCTCACACACACACACACACA

-320                                                                   ECO RI
CACACACACACACACACACACGTACACACGTACACACACACGTGCGCATGCTTGACGCATACGCGACGAATTCGAAAGCT

-240
CACCTCATTCCTCCCTCCTCACACCATCATCGGCATCCATAGAGACACGTGCGCGTAGAACGATACAGCTCACGCGCACA

-160
GAGAGGCGTGTTCTGGTCGTGCGCATCATCGGCACAGCACTGGCGGACGAAACTCGCACACAGGCACGCCGCCTCCTTTC

-80
ACCCGTCATAGATAGTTGAATTAGACGCCCTCCTCCTCCCTCATCATCGCCGTCGTCATCCGGGTCCGAGCACTACGAAG

+1
ATG
met
```

Figure 2. DNA sequence of the 5' end of the DHFR-TS gene. The DNA sequence from nucleotide -952 to the initiation codon (+1, met) is shown. The restriction sites for Sal I and Eco RI are shown above the sequence and the locations of the 5' ends of the DHFR-TS mRNAs as determined by S1 nuclease mapping are underlined. The two regions of homology with the *L. major* miniexon (J. Miller, unpublished data) are indicated by solid lines (a and b) located above the DNA sequence. The sequence of the miniexon found on the 5' end of the beta-tubulin gene is TGCTATATAAGTATCAGTTTCTGTACTTTAT (kindly provided by J. Miller).

Interestingly, the region from -348 to -301 contains a stretch of two polydinucleotides, $(CT)_4$ followed by $(CA)_{20}$. Other regions of the DHFR-TS gene contain polydinucleotides, such as poly-(GA) tracts in the 3' untranslated region of the DHFR-TS mRNA (27), and poly-(CA) tracts within the region flanking the 3' end of the DHFR-TS mRNA and many other regions in the vicinity of the DHFR-TS gene (J. Cordingley and SMB, unpublished data). Poly-(CA) tracts have also been found immediately upstream of the 5' ends of the mRNAs for beta-tubulin from *L. major* (J. Miller, unpublished data), alpha and beta tubulin from *L. enrietti* (45),

and a repetitive gene family from *T. cruzi* (48). Poly-(CA) has the capacity to adopt a Z-DNA configuration (49), has been shown to enhance gene expression in cultured mammalian cells (50) and appears in the vicinity of some genes which undergo rearrangement (51, 52; J. Cordingley and SMB, unpublished data). Whether this dinucleotide repeat plays any role in trypanosomatid gene expression is unknown. Small sequence repetitions other than dinucleotides can be found throughout the DHFR-TS region, as shown recently as well for the mini-exon repeat of *Trypanosoma brucei* (8), the origin and significance of which are unknown. Finally, a 24 amino acid open reading frame begins at position -273, which would be present in the longer mRNAs identified by S1 mapping.

In species other than trypanosomatids, "consensus" elements implicated in gene expression have been identified. These include the TATA box (53), CAAT box (54), Pribnow box (55), enhancers (56), Spl binding sites (57), etc. None of these elements appear within the 5' region of the DHFR-TS gene, nor is the consensus eukaryotic polyadenylation signal (58) present in the 3' end of the DHFR-TS mRNA (27).

Comparisons of the 5' regions of DHFR and TS genes of several species.

We next examined the 5' region of the DHFR-TS gene for homology with that of the DHFR genes from *E. coli* and several mammalian species, and that of the TS genes from *E. coli* (59) and mouse (21). Analysis of the DHFR genes from mouse (60), hamster (61) and human (62) has revealed the presence of a conserved sequence which functions as the DHFR promoter (63). This element, which can be repeated up to four times depending upon the species, can be further divided into a "GC" element and a "CAA" element (61). Dot matrix comparisons of the promoter region of these genes with the 5' region of the DHFR gene of *E. coli* (64) revealed that this species contains a previously unreported homology with the mammalian DHFR promoter region, occurring in *E. coli* at positions -134 to -87 with respect to the initiation codon. Table 1 shows the alignment of this *E. coli* element with the corresponding regions of the mammalian species. The homology of the *E. coli* element is 67% with the mouse DHFR consensus element

Table 1. Sequence comparisons of the DHFR promoter element from mammalian cells and homologous regions of _E. coli_ and _Leishmania_.

species	5' position[a]	homologous sequence	3' position[a]
mouse[b]	-282	TGCGC---GGCGGG-CCTTGGTGGGGGCGGrGCCTTAGCTGCACAA-ATAGxA	-132
hamster	-160	CT$^{C}_{T}$GT$^{G}_{T}$-GAGGCG$^{G}_{T}$GGCCTC$^{T}_{C}$GAT$^{GTT}_{TCA}$CAA-$^{A}_{G}$T$^{A}_{G}$GGA	-84
human	-133	TGCGCCGGGGCGGG---------GGGGCGGGGCCTCGCCTGCACAA-ATAGGG	-90
E. coli	-134	TGCG-----GCGAGTCCAGGGAGAGAGCGTGGACTCGCCAGCAGAATATAAAA	-87
L. major	-451	CCCCAGCTGCAGACAA	-436
consensus[c]		TGCG GCGrG GGGGCG$^{G}_{T}$GGCCTCxxCTGCACAA-ATAGxA	
		GC box CA box	

[a]relative to the initiation codon. 5' and 3' positions for the mouse and hamster demarcate the extreme borders of the multiple DHFR elements of these species. Gaps inserted for alignment of compared sequences are indicated with dashed lines. The letters r and x of the mouse and consensus sequences sequences designate purine and no consensus bases respectively among the compared sequences.
[b]the sequence shown is a consensus of the four mouse repeats.
[c]consensus is defined as three or more identities among the non-_Leishmania_ sequences.

(31/46), 57% with the hamster consensus element (20/35) and 74% with the human element (28/38). These values may be compared with the average per cent homology among the three mammalian elements, 85%. Surprisingly, the conservation between the _E. coli_ element and the mouse promoter element significantly exceeds the 44% homology observed between the protein coding regions of these genes (60; 64; SMB, unpublished data). The conservation of this element in _E. coli_ may be suggestive of a functional role, although it has not been experimentally tested. Several _E. coli_ genes have been shown to have regulatory sequences upstream of the consensus "-35" and Pribnow box elements (65,66).

Organismal and molecular evolutionary comparisons of ribosomal RNA, DHFR and TS suggest that _Leishmania_ and the Trypanosomatidae are more closely related to mammals and other eukaryotes than to prokaryotes (27,67). A simple prediction would be that elements conserved among _E. coli_ and mammalian DHFR genes ought to be present in the _Leishmania_ DHFR-TS gene as well, whose 5' end encodes the DHFR protein domain (27,28). However, we have been unable to find an element comparable to that shared by _E. coli_ and mammalian DHFR genes. The best homology is located at positions -436 to -451, (10-11/16 bp) which as shown in Table 1 would correspond to the "CAA" element of the other DHFR genes.

This degree of homology can be seen in comparisons among many other regions of the DHFR and TS genes. Thus, *Leishmania* appears to lack a sequence corresponding to the mammalian DHFR promoter as conserved as that found in *E. coli* . Finally, comparison of the 5' region of the *Leishmania* DHFR-TS gene with that of the mouse (21) and *E. coli* (59) TS genes did not reveal significant homology.

DISCUSSION

We have shown by S1 analysis that mRNA transcribed from the DHFR-TS gene displays heterogeneity at the 5' end consisting primarily of two species whose termini map to positions -245 and -375 relative to the start of translation (Fig. 1B). Two additional minor S1 products have been mapped to positions -195 and -295, with an additional minor product extending beyond nt -440. The 5' noncoding region of the *Leishmania* DHFR-TS mRNA is considerably longer than that of other DHFRs (23-115 nt; 60-62,64) and *E. coli* TS (22 nt; 59), but shorter than that of the TS of *Herpes samirii* (1207 nt; 68). We have not been able to detect any significant potential RNA secondary structure in the 5' non-coding region using the algorithm of the PC-Fold program of M. Zuker (69).

We find that the relative abundance of the major S1 products is not altered by 1) amplification of the DHFR-TS gene, 2) growth phase, nor 3) developmental stage. In this regard transcripts of the *Leishmania* DHFR-TS resemble those of the mouse DHFR, which maintain a constant 5' end profile as cells progress through the cell cycle (23). Figure 1C also shows that the level of the DHFR-TS mRNA is elevated in the R1000 line relative to the wild-type promastigotes and amastigotes, as expected for transcripts of an amplified gene. Further studies of the regulation of the DHFR-TS mRNA and protein, and the addition of mini-exon sequences, will appear elsewhere.

Comparisons of the 5' region of the *Leishmania* DHFR-TS gene with that of other trypanosomatid genes did not reveal any unambiguous evidence of conserved sequences, which in the absence of functional assays could be suggestive of important roles. This observation has been reported by many workers in their

analyses of genes isolated from *Leishmania* and other trypanosomatids (45, reviewed in 1). The 5' region of the *Leishmania* DHFR-TS gene further fails to exhibit homology with the conserved promoter of mammalian DHFR genes, an element which is shared by the *E. coli* DHFR gene. Several explanations may be formulated for these observations: potentially cis-acting regulatory signals may be remarkably cryptic in the Trypanosomatidae, or alternatively, these putative cis-acting elements may not be located in the DNA immediately upstream of the 5' end of the mature mRNA species. This latter possibility is supported by studies of several genetic loci in trypanosomes which suggest that transcription may possibly initiate as far as 40 kb 5' of the body of the mRNA (49, 70-72). Another hypothesis would be that the regulation of the DHFR-TS gene is different in *Leishmania* than in other species, accompanied by divergence of the 5' regulatory regions. Discriminating among these or other hypotheses would be greatly aided by the ability to perform genetic manipulations, a goal actively being pursued in many labs.

ACKNOWLEDGEMENTS

We thank Jacqueline Miller for extensive discussions and providing unpublished sequence data of the beta-tubulin and mini-exon of *L. major*, Tom Ellenberger for assistance in DNA sequencing, S. Sazer, and D. Dobson. This work was supported by NIH Grant AI 21903-02. GMK was supported by NIH Training Grant 5-T-32-GM07196 and KZ by Harvard University, Division of Medical Sciences. SMB is a Burroughs-Wellcome Scholar in Molecular Parasitology.

*To whom correspondence should be addressed

REFERENCES

1. Borst, P. (1986) Ann. Rev. Biochem. 55:701-32.
2. Benne, R. (1985) Trend Genet. 1:117-121.
3. Walder, J.A., Eder, P.S., Engman, D.M., Brentano, S.T., Walder, R.W., Knutzon, D.S., Dorfman, D.M. and Donelson, J.E. (1986) Science 233:569-71.
4. Cornelissen, A.W.C.A., Versperian, M.P., Toulme, J-J. Swinkels, B.W. and Borst, P. (1986) Nucleic Acids Res. 14: 5605-5614.

5. DeLange, T., Michels, P.A.M., Veerman, H.J.G., Cornellisen, A.W.C.A. and Borst, P. (1984) Nucleic Acids Res. 12: 3777-3790.
6. Parsons, M., Nelson, R.G., Watkens, K.P. and Agabian, N. (1984) Cell 38:309-316.
7. Nelson, R.G., Parsons, M., Barr, P.J., Stuart, K., Selkirk, M. and Agabian, N. (1983) Cell 34:901-909.
8. Campbell, D.A., Thornton, D.A. and Boothroyd, J.C. (1984) Nature 311:350-355.
9. Blakely R.L. (1969) in The Biochemistry of Folic Acid and Related Pteridines pp. 231-249. (North Holland, Amsterdam).
10. Bertino, J. and Johns, D. (1972) in Cancer Chemotherapy (I. Bradsky, ed.) vol. 2, pp. 9-22, (Grane and Stratton, N.Y.).
11. Kellems, R.E., Morhenn, V.B., Pfendt, E.A., Alt, F.W. and Schimke, R.T. (1979) J. Biol. Chem. 254:309-318.
12. Mariani, B.D., Slate, D.L. and Schimke, R.T. (1981) Proc. Nat. Acad. Sci. USA 78:4985-4989.
13. Jenh, C.-H., Rao, L. G. and Johnson, L.F. (1985) J. Cell Physiol. 122:149-154.
14. Navalgund, L.G., Rossana, C., Muench, A. and Johnson, L.F. (1980) J. Biol. Chem. 255:7386-7390.
15. Jehn, C-H., Geyer, P.K. and Johnson, L.F. (1985) Mol. Cell Biol. 5:2527-2532.
16. Ley, E.J. and Kellems, R.E. (1981) Mol. Cell. Biol. 1:961-971.
17. Farnham, P.J. and Schimke, R.T. (1985) J. Biol. Chem. 260: 7675-7680.
18. McGrogan, M., Simonsen, C.C., Smouse, D.T., Farnham, P.J. and Schimke, R.T. (1984) J. Biol. Chem. 260:2307-2314.
19. Gasser, C.S. and Schimke, R.T. (1986) J. Biol. Chem. 261: 6938-6946.
20. Sazer, S. and Schimke, R.T. (1986) J. Biol. Chem. 261:4685-4690.
21. Deng, T., Dawei, L., Jenh C-H., and Johnson, L. (1986) J. Biol. Chem. 261:16000-16005.
22. Setzer, D.R., McGrogan, M., and Schimke, R.T., (1982) J. Biol. Chem. 257:5143-5147.
23. Farnham, P.J. and Schimke, R.T. (1986). Mol. Cell. Biol. 6: 365-371.
24. Ferone, R. and Roland, S. (1980) Proc. Nat. Acad. Sci. USA 77:5802-5806.
25. Coderre, J.A., Beverley, S.M., Schimke, R.T. and Santi, D.V. (1983) Proc. Nat. Acad. Sci. USA 80:2132-2136.
26. Garrett, C.E., Coderre, J.A., Meek, T.D., Garvey, E.P., Claman, D.M., Beverley, S.M. and Santi, D.V. (1984) Mol. Biochem. Parasitol. 11:257-265.
27. Beverley, S.M., Ellenberger, T.E. and Cordingley, J.S. (1986) Proc. Nat. Acad. Sci. USA 83:2584-2588.
28. Grumont, R., Washstein, W.L., Caput, D. and Santi, D.V. (1986) Proc. Nat. Acad. Sci. U.S.A. 83:5387-5391.
29. Beverley, S.M., Coderre, J.A., Santi, D,V. and Schimke, R.T. (1984) Cell 38:431-439.
30. Ebrahimzadeh, A. and Jones, T.C. (1983) Am. J. Trop. Med. Hyg. 32:692-702.
31. Beverley, S.M., Ismach, R.B. and McMahon Pratt, D. (1986) Proc. Nat. Acad. Sci. USA, 84:484-488.

32. Jaffe, C.L., Grimaldi, G. and McMahon Pratt, D. (1984) in Genes and Antigens of Parasites, 2nd edn. C. Morel, ed. pp 47-91, (Rio de Janeiro: Fundacao Oswaldo Cruz).
33. Cathala, G., Savouret, J.F., Mendez, B., West, B., Karin, M., Martial, J.A. and Baxter, J.D. (1983) DNA 2:329-335.
34. Aviv, H. and Leder, P. (1972) Proc. Nat. Acad. Sci. USA 69:1408-1412.
35. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor Laboratory, New York).
36. Feinberg, A.P. and Vogelstein, B. (1983) Anal. Biochem. 132: 6-13.
37. Berk, A.J. and Sharp, P.A. (1977). Cell 12:721-732.
38. Messing, J. (1983). Meth. Enz. 101:20-78.
39. Maxam, A.M. and Gilbert, W. (1977) Proc. Nat. Acad. Sci. USA 74:560-64.
40. Sanger, F., Nicklen, S. and Coulson, A.R. (1977). Proc. Nat. Acad. Sci. USA 74:5463-5467.
41. Biggins, M.D., Gibson, T.J. and Hong, G.F. (1983). Proc. Nat. Acad. Sci. USA 80:3963-3965.
42. Van der Ploeg, L.H.T., Liu, A.Y.C., Michels, P.A.M. Delange, T. and Borst, P. (1982) Nucleic Acids Res. 10: 3591-3604.
43. Sather, S. and Agabian, N. (1985) Proc. Nat. Acad. Sci. USA 82:5695-5699.
44. Tschudi, C., Young, A.S., Ruben, L., Patton, C.L. and Richards, F.F. (1985) Proc. Nat. Acad. Sci. USA. 82, 3998-4002.
45. Landfear, S.M., Miller, S.I. and Wirth, D.F. (1986) Mol. Biochem. Parasit. (in press).
46. Mount, S.M. (1982) Nucleic Acids Res. 10:459-472.
47. Reed, R. and Maniatis, T. (1985) Cell 41:95-105.
48. Gonzalez, A., Lerner, T.J., Huecas, M., Sosa-Pineda, B., Nogueira, N. and Lizardi, P. M. (1985) Nucleic Acids Res. 13:5789-5804.
49. Johnston, B.H. and Rich, A. (1986) Cell:713-724.
50. Hamada, H., Seidman, M., Howard, B.H. and Gorman, C.M. (1984) Mol. Cell. Biol. 4:2622-2630.
51. Liu, A.Y.C., Van der Ploeg, L.H.T., Rijsewijk, F.A.M. and Borst, P. (1983) J. Mol. Biol. 167:57-75.
52. Nishioka, Y. and Leder, P. (1980) J. Biol. Chem. 255:3691-3694.
53. Breathnach, R. and Chambon, P. (1981) Ann. Rev. Biochem. 50: 349-383.
54. Efestratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.W., O'Connell, C.O., Spritz, R.A., DeReil, J.K., Forget, B.G., Weissman, S.M., Slighton, J.L., Blechl, A.E., Smithes, O., Baralle, F.E., Shoulders, C.C. and Proudfoot, N.A. (1980) Cell 21:653-668.
55. Pribnow, D. (1975) Proc. Nat. Acad. Sci. USA 72:784-788.
56. Banerji, J., Rusconi, S. and Schaffner, W. (1981) Cell 27:299-308.
57. Dynan, W.S. and Tjian, R. (1983) Cell 35:79-87.
58. Proudfoot, N.J. and Brownlee, G.G. (1976) Nature 263:211-214.
59. Belfort, M., Maley, G., Pedersen-Lane, J. and Maley, F. (1983) Proc. Nat. Acad. Sci. USA 80:4914-4918.

60. Crouse, G.F., Simonsen, C.C., McEwan, R.N. and Schimke R.T. (1982) J. Biol. Chem. 257:7887-7897.
61. Mitchell, P.J., Carothers, A.M., Han, J.H., Harding, J.D., Kas, E., Venolia, L. and Chasin, L.A. (1986) Mol. Cell. Biol. 6:425-440.
62. Masters, J.N., and Attardi, G. (1986) Mol. Cell. Biol. 5: 493-500.
63. Farnham, P. and Schimke, R.T. (1986) Mol. Cell. Biol. 6: 2392-2401.
64. Smith, D.R. and Calvo, J.M. (1980). Nucleic Acids Res. 8: 2255-2272.
65. Dunn, T., Hahn, S., Ogden, S. and Schleif, R. (1984) Proc. Nat. Acad. Sci. USA 81:5017-5020.
66. Irani, M.H., Orosz, L. and Adhya, S. (1983) Cell 32:783-788.
67. Gray, M.W. (1981) Mol. Cell. Biol. 1:347-357.
68. Bodemer, W., Niller, H.H., Nitsche, N., Scholz, B. and Fleckenstein, B. (1986) J. Virol. 60:114-123.
69. Salser, W. (1977) Cold Spring Harbor Symp. Quant. Biol. 42: 985-1002.
70. Bernards, A., Kooter, J.M. and Borst, P. (1985) Mol. Cell. Biol. (5):545-553.
71. Van der Ploeg, L.H.T., Shea, C., Polvere, R. and Gwo-Shu, M.L. (1986) in _Molecular Strategies of Parasite Invasion, UCLA Symposia on Molecular and Cellular Biology_, new _series vol 2._ _eds. N. Agabian, H. Goodman_ and _N. Nogueira._ (Alan R. Liss. Inc. N.Y. in press).
72. Swinkels, B.W., Gibson, W.C., Osinga, K.A., Kramer, R., Veeneman, G.H., Van Boom, J.H. and Borst, P. (1986) EMBO J. 5:1291-1298.