Characterization of a differentially expressed histone H2B Leishmania enriettii: a thesis.

by

Jane E. Genske

Oregon Health Sciences University

School of Medicine

May 1990

APPROVED:

[redacted]

(Professor in Charge of Thesis)

[redacted]

(Chairman, Graduate Council)

TABLE OF CONTENTS

List of Figures	v
Dedication	vi
Acknowledgements	vii
Abstract	viii
Introduction	1
The Organism	1
The Disease	2
Molecular Biology of *Leishmania*	3
Project Background	4
Histones	6
Materials and Methods	10
Results	15
Sequence Analysis of cDNAs and Homology Search	15
Structure of cDNAs	16
Structure of H2B Genes	17
Densitometry	18
Chromosomal Analysis	18
DNA Synthesis Inhibition and Histone H2B mRNA Levels	19
S1 Mapping Experiments; Possible Precursor Transcripts	21
Discussion	23
Developmentally Regulated Genes in *Leishmania*	23
Identification of a *Leishmania* Histone H2B	23

Structure of *Leishmania* Histone H2B mRNAs 24

Genomic Arrangement of Histone H2B Genes 26

Regulation of Histone H2B Expression 27

Possible Polycistronic Transcripts of Histone H2B Genes 29

Homology to Eukaryotic 3' Processing Signal 30

Future Directions 32

Summary 32

References 34

LIST OF FIGURES

1. Life cycle 41
2. Northern blot of RNA from two life stages 42
3. Genomic Southern blot 43
4. cDNA alignment 44
5. Protein alignment 45
6. Primer extension analysis 46
7. Stem-loop structures 47
8. Northern blot of poly A+ RNA 48
9. Restriction map of genomic clone P6.4 49
10. PS3.5 sequence data 50
11. Densitometry analysis 51
12. Chromosomal analysis 52
13. ^{3}H-Hypoxanthine incorporation in DNA synthesis 53
14. Northerns of drug experiments 54
15. S1 mapping of P6.4 55
16. Schematic of mRNA structures 56

To my parents, Thomas and Mary Genske, who have supported and encouraged me in everything I have done.

And to Michael Widmer, my friend and soon-to-be husband, who has been a source of support, encouragement, love and happiness at the times I have needed it most.

Acknowledgements

I would like to acknowledge Alison Gammie, who has not only jumped through the hoops with me, but has also been a wonderful friend and source of strength in my graduate years, this is the final carrot.

I would like to acknowledge the members of my committee for continued advice and encouragement throughout this work. Most importantly, I would like to thank Dr. Scott Landfear, who introduced me to the wonderful world of parasitology, and who taught me molecular biology and the patience necessary to enjoy it.

ABSTRACT

The disease leishmaniasis afflicts millions of people each year, yet relatively little is known about the molecular biology of the causative organism. *Leishmania* spp. colonize both an insect vector and a vertebrate host during their life cycle. In the process of the life cycle, the parasite undergoes morphological and biochemical transformations necessary to adapt to the two very different environments of the insect gut and the vertebrate macrophage. It is reasonable to expect that there are differentially expressed genes which are involved in this adaptation. The purpose of this work was to characterize a previously cloned developmentally regulated gene from *L. enriettii*. It was determined by sequence analysis of cDNAs and genomic DNA that this regulated gene encodes a histone H2B. Comparison of this H2B sequence with other known H2Bs, confirmed that this is the most divergent histone H2B identified so far. The structure of the *Leishmania* H2B mRNA is unique in containing *both* a poly-A tail and a novel hairpin loop. The genomic DNA sequence shows that the H2B genes are arranged in tandem repeats and that the coding regions are conserved. Southern blots and chromosome gels identify three clusters of H2B genes dispersed on two different chromosomes. Unlike other eukaryotes, the levels of H2B mRNA are not decreased by DNA synthesis inhibitors, suggesting an alternative mechanism of regulation. S1 mapping experiments have yielded preliminary evidence consistent with the tandemly arrayed H2B genes being transcribed as polycistronic transcripts.

INTRODUCTION

The Organism

Leishmania are parasitic protozoa belonging to the order Kinetoplastidae, family Trypanosomatidae. Members of this family have a single flagellum, a nucleus, and a single mitochondrion called the kinetoplast. Species belonging to this family are referred to as trypanosomes or haemoflagellates, and are among the most primitive eukaryotes known; they are placed well below yeast on a phylogenetic tree (1). *Leishmania* species require two hosts in order to complete their life cycle, a vertebrate and an insect. The insects serving as vectors of disease, belong to either the genus *Phlebotomus* or *Lutzomyia*, and are more commonly referred to as sandflies. The cycle (figure 1) begins when the infected female sandfly takes a blood meal from a mammalian host. Infectious parasites, the promastigotes, are released into the blood and are quickly taken up by macrophages in the skin, mucosa or viscera. Once inside the macrophage, the promastigotes undergo a morphological transformation into amastigotes. Amastigotes are ovoid in shape, lack flagella and are obligate intra-macrophage parasites. They exist in the lysosome, where they are resistant to lysosomal enzymes. Amastigotes replicate *via* binary fission to the point where they cause rupture of the host macrophage. Upon release into the bloodstream they are quickly internalized, again by macrophages, and the cycle continues. When another sandfly takes a blood meal from an infected mammal, amastigote-containing macrophages are also ingested. These will lyse in the gut of the sandfly, and another morphological transformation occurs in which the amastigotes differentiate into the promastigotes, which are long, slender, flagellated and extracellular organisms. These promastigotes live in the gut of the sandfly and begin migrating to the salivary glands upon maturation. The cycle is completed

when mature, infectious promastigotes are injected into the skin of a mammalian host (2).

The Disease

Leishmaniasis is one of the most important infectious diseases in the tropics. There are an estimated 12 million new cases of leishmaniasis diagnosed each year (3). In the Old World, the disease is endemic around the Mediterranean Sea; in North and East Africa; in the Near East, southern Russia, and the Middle East; and in Afghanistan, India, and China. In the New World, leishmaniasis is found from the U.S.-Mexico border through Central America to South American countries such as Colombia, Ecuador, Venezuela, Peru, and Brazil (3).

There are three types of diseases caused by *Leishmania*: cutaneous, mucocutaneous, and visceral. Each is caused by different species of parasites. Biologic and biochemical criteria suggest that the *Leishmania tropica* complex of organisms causes cutaneous disease in the Old World, the *Leishmania mexicana* and *Leishmania braziliensis* complexes cause cutaneous and mucosal disease in the New World, and *Leishmania donovani* complexes cause visceral disease (3).

Cutaneous leishmaniasis characteristically occurs on an exposed area of the body and begins as a red papule; enlarges to form a shallow ulcer with raised, red margins; and then heals with scarring. The progression of disease to healing generally takes a year.

Mucocutaneous disease involving lymph nodes, usually consists of multiple lesions. *L. braziliensis* may also cause lesions of the oronasal mucosa. Mucosal involvement appears to be a late, metastatic complication of cutaneous leishmaniasis, with onset 2-25 years after healing of a primary cutaneous

lesion. The involvement of the oronasal mucosa leads to massive tissue destruction resulting in secondary infections which can be fatal.

Visceral leishmaniasis (kala-azar) is characterized by fever, hepatosplenomegaly, hypergammaglobulinemia and pancytopenia. Usually, death results in untreated cases due to secondary infections.

The current treatment for mucosal and visceral leishmaniasis involves pentavalent antimonials. These drugs are expensive and have serious side effects. There is a need to develop better (less toxic, less expensive) drugs. There is no vaccine available at the present time, though development of a practical vaccine is actively being studied.

Molecular Biology of Leishmania

Trypanosomatids display many unusual features of gene expression. They contain a single mitochondrion or kinetoplast, which has maxi-circle (about 20 kb) and mini-circle DNA (about 1 kb). Maxi-circle DNA encodes mitochondrial genes, but the function of mini-circle DNA remains unknown. It has been demonstrated that messages transcribed from kinetoplast DNA can undergo a transcriptional editing process, in which additional ribonucleotides, not encoded by the DNA, are inserted into the mRNA (4,5). Recently, several small RNAs encoded in intergenic regions of the maxi-circle DNA of *L. tarentolae*, have been identified, and have been linked with the transcriptional editing process (52). These small RNAs, called guide RNAs or gRNAs, contain sequences that represent precise complementary versions of the edited regions of mature mRNAs. It is hypothesized that gRNAs are able to hybridize to kinetoplast mRNAs, both upstream and downstream of the point of editing, and guide a number of cleavage reactions, uridylate additions or deletions, and religation to cause the change in mRNA sequence.

Another unusual feature is that all cytoplasmic mRNAs examined to date, lack intervening sequences and contain a 39 basepair (bp) spliced leader at the 5' end of the message (6,7). The leader sequence is added onto the 5' end in a reaction similar to cis-splicing, which occurs in the removal of intervening sequences in eukaryote mRNAs, in that there are typical consensus splice acceptor and splice donor sequences, and that a branched intermediate is involved in the reaction. It is unique, however, in that two separate mRNAs are spliced together. The function of this spliced leader sequence is not known.

Finally, many of the genes identified so far are arranged in tandem repeats, such as the tubulin genes and a transporter gene recently identified in our laboratory (8,9). There is evidence in the African Trypanosomes that tandemly arrayed genes may be transcribed as large units to form polycistronic transcripts. This has been demonstrated using heat-shock to block processing of putative precursors (10,11) and UV transcription mapping to determine the size of transcription units (12).

It should also be mentioned that no consensus sequences for promoters, polyadenylation signals, ribosome-binding or other regulatory sites have been identified. Thus many aspects of gene expression still remain to be elucidated.

Project Background

The original impetus for this research was to identify genes which are developmentally regulated during the *Leishmania* life cycle. Identifying genes which are preferentially expressed in one stage should give insight into the organism's ability to adapt to two very different environments. Obvious environmental differences between the sandfly gut and the macrophage lysosome include pH, temperature, and nutrient availability. The pH of the

lysosomal vacuole is approximately 4.6, whereas the pH of the gut of the sand fly is closer to 7. The body temperature of mammals is 37°C compared to 25°C for the sandfly. Though the precise composition of nutrients available either in the lysosome or the insect gut is not known, it is reasonable to assume they are different. It is also known that the parasite itself differs, at least morphologically, between the two host stages.

For these studies, we have used *Leishmania enriettii*, a parasite which is non-pathogenic to humans, and therefore safer for laboratory work. Both life cycle stages can be studied in the laboratory: promastigotes can be grown in tissue culture medium at 26°C, and amastigotes can be grown in Guinea pigs and isolated from the tissue lesions (58). The basic biology of this parasite is very similar to that of the human pathogens.

In order to identify genes which were specific to the promastigote stage, a cDNA library (45) for the promastigote stage was constructed using the vector λgt11 (55). A differential hybridization using radiolabelled cDNA probe templated from either promastigote or amastigote polyadenylated RNA was performed to identify genes which hybridized to only radiolabelled promastigote cDNA, suggesting these genes are preferentially expressed in the promastigote stage. One such clone, designated c61-1, was chosen for this study. A Northern blot of promastigote and amastigote RNA, was probed with the cDNA c61-1, and a 750 bp message was identified, which was much more abundant in the promastigote stage than the amastigote stage (figure 2). The blot was re-hybridized with radiolabelled ribosomal RNA (56) to demonstrate equal amounts of RNA were loaded in each lane. This result indicates that this gene is differentially expressed during the parasite life cycle.

To study the arrangement of this gene in the *Leishmania* genome, restriction digests of genomic DNA were performed and Southern blotted.

Multiple fragments were observed that hybridized to the cDNA c61.1, suggesting multiple genes. The cDNA was also used to probe *Pst* I digested genomic DNA from promastigotes. Three fragments hybridized to the cDNA (figure 3), and the middle fragment, (fragment b, 6.4 kilobase pair (kb) in size) was cloned into the plasmid pBluescript SK+.

One of the primary objectives of the research reported here was to determine the function of this gene. A reasonable approach to this problem was to sequence the gene and determine whether it encoded a protein that was related in sequence to another protein of known function. This approach had previously been used in our laboratory to identify a gene encoding a developmentally regulated membrane transport protein (9). From sequence analysis of cDNAs, which will be discussed later, it was determined that the differentially expressed gene studied here encodes a histone H2B.

Histones

Histones are a class of basic proteins that associate with each other and with nuclear DNA to form the nucleosome, the fundamental unit of chromatin structure (14). There are four core histone types-H2A, H2B, H3, and H4, which are organized into an octamer, containing two copies of each type, and which form the body of the nucleosome. The fifth histone type, H1, is present in half the amount of each core histone and is located in linker regions of the chromatin, between individual nucleosome cores. Each class of histone is characterized by the relative amount of lysine and arginine residues present in each polypeptide; H2B is generally slightly lysine rich. The role of histones in gene regulation remains controversial, though it is widely accepted that they are fundamental in organizing and compacting DNA.

The structures of histones in general are highly conserved between diverse animal phyla and even between the animal and plant kingdoms, with

H3 and H4 being the most conserved (14). The typical H2B mRNA is 400-700 bp, it encodes a 14-17 kilodalton (kd) protein of approximately 120 amino acids. The genomic arrangement of histone genes is also conserved. All five histone genes are generally clustered in the genome, and commonly in tandem repeats varying in copy number from 150 up to 700 depending on the species (15). However, the yeast *Saccharomyces cerevisiae* contains only two copies of each histone gene (57). These are arranged in duplicate pairs (H2A coupled to H2B and H3 coupled to H4), but no pair is closely linked to any other. Heterogeneous isotypes are more common in the histones H2A, H2B, and H1, and have been shown to differ not only in nucleic acid, but also in amino acid sequence. In some organisms different isotypes are expressed at different times during embryonic development (13).

The production of histones is primarily necessary when the cell is undergoing DNA synthesis in S phase; at this time, new histones are required to package the newly synthesized DNA. In most systems studied, the synthesis of new histone proteins and the accumulation of histone mRNAs occurs primarily during S phase (16,17,18,19). The coupling of histone synthesis to DNA replication is the result of multiple levels of regulation, and this is an extensively studied area of histone gene expression.

The rate of histone gene transcription in both yeast and higher eukaryotes increases approximately three- to five-fold at the end of G1, as cells enter S phase (20,21). The sequences involved in regulating transcription are adjacent to the histone promoter. Reporter genes, such as neomycin resistance, fused to the 5' sequences of an H3 gene from hamster, are transcribed in a cell-cycle dependent manner, similar to histone genes (22).

The stability of histone mRNAs is also a major determining factor of histone levels. In higher eukaryotes, the half-life of many histone mRNAs

fluctuates from 45-60 minutes during S phase, to 10-15 minutes at the end of S phase or in the presence of DNA synthesis inhibitors (20,23). In higher eukaryotes, histone message degradation is dependent on a conserved stem-loop at the 3' end. If the 3' histone stem-loop is fused to the 3' end of a normally stable, constitutively expressed gene, such as an α-globin, the half life of this mRNA will be regulated during the cell cycle in a way which is similar to the regulation of histone mRNA half lives (24). Furthermore, based on the following results, it is believed that translation of the histone mRNA is required for its increased rate of degradation outside of S phase: (1) histone mRNAs are stable in the presence of protein synthesis inhibitors (25), (2) histone mRNAs with a termination codon inserted early within the protein coding region are stable during inhibition of DNA synthesis (26), (3) the stem-loop is required to be within a certain distance (approximately 300 bp) of the termination codon for degradation to occur (26), and (4) degradation occurs 3' to 5' (27). One possible explanation for these results would be a ribosome-associated nuclease which is activated when it comes into contact with the 3' stem-loop. Transcriptional regulation and message stability together account for a 30- to 50-fold change in the observed level of histone mRNA levels during the cell cycle.

The final step of regulation is post-transcriptional and is the processing of the 3' end to the mature message. It has recently been shown that extracts from cells in G1 phase lack the 3' processing activity (28). The molecular basis of the change in efficiency of 3' end formation is not known.

There is a second class of histone genes, the variant or basal histones, whose expression is not regulated during the cell-cycle; rather they are synthesized at a basal level throughout the cycle. Unlike cell cycle regulated histone mRNAs, they are polyadenylated, they lack a hairpin structure at the

3' end, and they contain intervening sequences. Their role in chromatin packaging appears to be in replacement of cell-cycle regulated histones, though more research is needed before classifying their function.

Yeast histone mRNAs are different from higher eukaryote mRNAs in that they lack the conserved hairpin structure at the 3' end, and they are poly-adenylated. However, all yeast histone mRNAs identified to date are cell-cycle regulated. Although they do not have the hairpin loop, the 3' end of the message is still required for proper regulation, as determined by deletion analysis (29).

In this thesis, I have studied the H2B genes and mRNAs from *Leishmania enriettii*. Several unusual features of their structure and regulation have been identified which distinguish these histone genes from those of either yeast or the higher eukaryotes.

MATERIALS AND METHODS

Tissue culture and nucleic acid

Promastigotes of *L.enriettii* were grown and maintained at 26°C in DME-L, containing 10% fetal bovine serum, 100 μM xanthine and 0.0005% hemin, as previously described (30). Cells were harvested at concentrations between 2 and 5 x 10^6 cells per ml. Cells were counted using either a hemocytometer or a Coulter counter. Genomic DNA and total RNA were isolated by phenol/chloroform extraction, as described previously (31,32). Southern and Northern blotting were performed as described (31,33). Radiolabelling of plasmid clone inserts was done by the random oligonucleotide priming method (34). Hybridizations of blots were performed at 42°C for at least 12 hr in 50% formamide, 5x SSC (1x SSC is 0.15 M NaCl, 0.015 M Na citrate, pH 7.0), 10x Denhardt's solution, 20 mM sodium phosphate pH6.5, 0.1% SDS, 250 μg/ml denatured herring sperm DNA. Filters were washed twice at room temperature and then three times at 55°C in 0.1x SSC, 0.1% SDS, 5 mM EDTA.

DNA sequencing and deletions

Double stranded DNA was sequenced using the enzymatic method of Sanger et al., (35), using ^{35}S-dATP (New England Nuclear, Boston, MA) and the Sequenase kit (United States Biochemical Corp., Cleveland, OH). Nested deletions of cloned DNA were obtained by the exonuclease III method (36) using the Erase-a-Base kit (Promega Biotech, Madison, WI) as recommended by the manufacturer. Extension products were separated on 6% acrylamide/8M urea gels (37). Both strands of the 3,524 bp PS3.5 genomic clone, and of the three cDNAs, c18, c37, and c43, were sequenced using double-stranded plasmid DNA as template (38). Plasmid DNA was prepared by alkaline lysis of cells and PEG precipitation as described (39).

Three synthetic oligos were purchased from Research Genetics, Birmingham, AL., which were used for sequencing the three cDNAs:

Oligonucleotide A: 5'-CTC-GTA-CGT-GAA-CGA-3'

Oligonucleotide B: 5'-GGT-GGG-CCG-CGG-TGC-3'

Oligonucleotide C: 5'-GCT-TCG-GCT-TGC-GGT-GCG-AC-3'

Computer assisted data analysis

The homology search of the NBRF protein database employed the FASTA algorithm (47) and alignment of protein sequences the BESTFIT algorithm (46). The stability of RNA stem-loops was calculated using the Stemloop algorithm. All algorithms were performed on a Micro VAXII using the software provided by the Wisconsin Genetics Computer Group (40). I would like to thank Michael Collard and Kenneth Harper for running these programs.

Primer extension analysis

8.0 pmoles of oligonucleotide primer C were end-labelled with γ^{32}P-ATP and T4 polynucleotide kinase (41). 0.4 pmoles of end-labelled primer was hybridized to 3 µg of promastigote polyadenylated RNA in hybridization buffer (0.01 M Pipes, pH 6.4, 0.40 M NaCl) for 2 hr at 65°C. Following incubation the hybrid was ethanol precipitated and resuspended in 25 µl of buffer containing 50 mM Tris-HCl (pH 8.3), 140 mM KCl, 7 mM $MgCl_2$, 10 mM dithiothreitol, 0.06 units RNasin, and 100 µM of each deoxynucleotide. Five units of avian myoblastosis virus reverse transcriptase (Life Sciences, St. Petersbrug, FL) was added and the reaction was incubated at 42°C for 30 minutes. Extension products were phenol/chloroform extracted and ethanol precipitated and resuspended in loading buffer, containing 95% formamide, 20 mM EDTA 0.05% Bromophenol Blue, and 0.05% Xylene Cyanol FF. Primer extension products

were compared with sequence from genomic DNA primed with the same oligonucleotide, using a 6% acrylamide/8 M urea sequencing gel.

Densitometry analysis

The copy number of H2B genes was estimated using densitometry. Southern blots were exposed to pre-fogged Kodak XAR-5 film in the presence of a Dupont Cronex Lightning Plus intensifying screen. Autoradiograms were scanned with a Bio Rad model 620 video densitometer, and the optical densities underneath peaks were automatically integrated using the Bio Rad 1-D Analyst software and an IBM AT personal computer.

Pulsed field gel electrophoresis (PFGE)

L. enriettii chromosomes were prepared for analysis by the following procedure. 2 x 10^8 cells were pelleted and washed with phosphate buffered saline (PBS), resuspended in 50 µl PBS and mixed with an equal volume of 1% In Cert agarose (FMC Bio Products, Rockland, ME) and poured into a mold, yielding 2 x 10^7 cells/block. Blocks were incubated for 48 hours at 55°C in 0.5 M EDTA pH 8.0, 1% sarkosyl, 2 mg/ml proteinase K, rinsed with 10 mM Tris pH 8.0, 1 mM EDTA and stored at 4°C. Electrophoresis was performed using a rotating gel apparatus (42) designed by Mr. Steven Smith (Los Lunas, NM). Gels were run for 30 hours, at 15°C, at 5 volts/cm with a 60 second dwell time, and a 120° rotation angle. Gels were then stained with ethidium bromide (1 µg/ml) and photographed and then incubated in 0.25 N HCl for 30 minutes followed by two 30 minute incubations in 0.5 N NaOH, 1.5 M NaCl (denaturing solution). Gels were blotted onto nitrocellulose filters using denaturing solution for the transfer. Yeast chromosomal size markers (43) were prepared in a similar manner from strain A364a of *Saccharomyces cerevisiae*.

S1 nuclease mapping

S1 mapping was performed as described (33). The appropriate annealing temperature was determined to be 62°C by incubating multiple samples of linearized double stranded DNA, at different temperatures, and then treating with S1 nuclease, which would be able to degrade single-stranded or denatured DNA. Sensitivity to S1 nuclease was monitored by running the samples on a 1.5% alkaline agarose gel (41), which contained 1 mM EDTA and 50 mM NaCl, with a running buffer of 30 mM NaOH and 1 mM EDTA.

Mapping experiments were performed using 30 µg of total promastigote RNA and 200 ng of linearized P6.4 DNA. DNA and RNA were denatured in 80% formamide, 0.4 M NaCl, 40 mM pipes buffer, pH 6.4, and 1 mM EDTA, at 72°C for 10 minutes, and then annealed at 62°C for 3 hours. Samples were then diluted into ice cold S1 buffer, containing 30 mM NaAcetate, pH 4.6, 250 mM NaCl, 5% glycerol, 1 mM $ZnSO_4$, and 20 µg/ml Herring sperm DNA and snap frozen in dry ice-ethanol. S1 nuclease (300-500 units) was added to thawed samples, and incubated at 45°C for 30 minutes. Following phenol/chloroform extractions and ethanol precipitation, samples were analyzed on alkaline agarose gels as described above. Gels were blotted according to the Southern blotting procedure and hybridized with radiolabelled c18 probe.

Inhibition of DNA synthesis

L. enriettii promastigotes were grown in culture as described, except the media contained 10 µM hypoxanthine in place of xanthine, as a purine source. Measurement of DNA synthesis was the same as previously described (44). Basically, 40 µCi of ^{3}H-hypoxanthine (gift from Dr. Buddy Ullman, OHSU) was added to 50 ml of exponentially growing cultures at a density of 2-4 x 10^6 cells/ml. These cultures were then used to monitor DNA synthesis, by measuring incorporation of ^{3}H-hypoxanthine. Two ml samples were harvested

in triplicate at two hour intervals from 0 to 8 hours. Cell number was determined using a hemocytometer. Cells were first pelleted and then washed with PBS, and incubated in 1 ml of 0.3 N KOH at 37°C overnight. Aliquots (50 µl) were transferred to filters which were pre-treated with 5% TCA- 1% sodium pyrophosphate (TCA-PPi) and dried. Filters were washed 4 times in TCA-PPi at 4°C and then rinsed with 95% ethanol. Filters were dried and counted in OMNIFLUOR liquid scintillation fluor (NEN Research Products, Boston MA) using a Beckman LS 3800 scintillation counter.

For measurement of inhibition of DNA synthesis, the procedure was the same as above, except that either 10 µM aphidicolin (Sigma, St. Louis, MO) or 5 mM hydroxyurea (Sigma, St. Louis, MO) was added to cells at time zero.

To quantitate the effect of DNA synthesis inhibition on H2B mRNA levels, cells were grown to a similar density, exposed to either 10 µM aphidicolin or 5 mM hydroxyurea, and RNA was isolated at t=0 hr, t=8 hr, t=10 hr, and t=12 hr. Northern blots containing 5 µg of total RNA from each sample, were probed with the c18 (H2B) cDNA, the Pro-1 transporter clone (9), and with the β-tubulin clone, pLEβ3 (33).

RESULTS

Sequence Analysis of cDNAs and Homology Search

The original cDNA clone c61-1, is 296 bp in length, as determined by dideoxy-nucleotide sequencing (data not shown). To determine the structure of this developmentally regulated gene, it was necessary to obtain a longer cDNA. For this a second cDNA library was constructed by Dr. Scott Landfear using the vector λgt10 (55). Three cDNAs which hybridized to the c61-1 insert were isolated from this library. They were designated c18, c37, and c43 and each contains an insert of approximately 600 bp. Inserts were subsequently subcloned into the plasmid pBluescript SK+. All three were sequenced (figure 4) in both directions using double stranded DNA obtained from mini-preps as template. The sequence of c37 was identical to that of c18, except that c37 was 107 base pairs shorter than c18. In addition to the primers supplied in the Sequenase kit, two synthetic 15 bp oligonucleotide primers (oligonucleotides A and B above) were purchased, which hybridize to the sequences overlined in figure 4.

Computer analysis of these cDNA sequences identified an open reading frame encoding 111 amino acids followed by an ochre stop codon (figure 10). A search of the NBRF protein databank, using the above deduced sequence, identified 23 homologous protein sequences, all of which were histone H2Bs. The BESTFIT algorithm was used to align the deduced protein sequence with the embryonic histone H2B (1) from sand sea urchin (figure 5), species *Psammechinus miliaris* (53), with an identity of 48%. A single gap at the amino terminus of Pro-2, is necessary for maximal alignment. From the alignment, it is clear that the highest amount of amino acid conservation occurs at the carboxy terminus of the sequence. This is true for all H2Bs identified so far (54).

Structure of cDNAs

The two cDNAs, c18 and c43 are close to full length. The nucleic acid sequence of c18 and c43 are identical in the coding region; however, the sequences begin to diverge in the 3' untranslated region (figure 4). The sequence of c37 is truncated at the 5' end, and begins at nucleotide 107 of the cDNA c18. The 3' untranslated region is identical to that of c18, except it has a 57 bp poly-A tail. It was determined that c18 and c43 contain sequence upstream of the start codon and that all three cDNAs are polyadenylated.

The 5' end of the H2B mRNA was mapped by primer extension using oligonucleotide C as primer. To determine the length of these primer extension products, a sequencing ladder was run along with the primer extension products (figure 6), using a subclone of the H2B genomic repeat (see below) as template and oligonucleotide C as primer. Several extension products were identified (figure 6), all approximately 85-90 nucleotides upstream from the 5' end of the c18 insert. Multiple products may result from multiple mRNAs with heterogenous sequence at the 5' end, since the histone mRNAs are encoded by a family of genes. In addition, multiple extension products could arise from the heavy methylation (59) on the first four residues of the spliced leader which is attached to all *Leishmania* RNAs (31); these methylated bases prevent complete progress of the reverse transcriptase through the first four bases of the mRNA (59).

It is known that cell-cycle regulated histones from higher eukaryotes contain a conserved 6 bp stem and a 4 bp loop at the 3' end of mRNAs (70). This stem-loop causes the histone mRNAs to become destabilized outside of the S phase of the cell cycle or in the presence of DNA synthesis inhibitors. I have identified a 7 bp stem, with a single mismatch, and a 4 bp loop at the 3' end of each cDNA (figure 7). The stem-loop in the c43 cDNA is quite different in

sequence from the stem-loop that occurs in both the c18 and c37 cDNAs (c37 not shown, it is identical with c18, figure 7). There is little similarity between either of the two *Leishmania* stem-loops and the consensus stem-loop sequence observed in other eukaryotes (61,70).

It was known from sequence data (figure 4) that the cDNAs isolated represented polyadenylated RNAs. To determine whether or not all of the H2B copies contained poly A tails, David Stein in our laboratory probed Northern blots containing total promastigote RNA, poly A+ RNA, and poly A- RNA with the c18 cDNA clone and with a control probe for another mRNA encoding the Pro-1 transporter (9). Almost all hybridization to the c18 probe occurred in the lane containing either total or poly A+ RNA (figure 8), with very little hybridization in the poly A- lane. This result indicates that most or all of the histone H2B mRNAs are polyadenylated.

Structure of H2B genes

To study the structure of the H2B genes, I have mapped and sequenced a genomic clone which contains multiple copies of the H2B sequence. Genomic fragments containing H2B sequence were previously identified by hybridizing blots of restriction digested genomic DNA with the c61-1 insert (figure 3). One of these fragments (band b, figure 3) was cloned by ligating a size selected fraction of *Pst* I digested genomic DNA into *Pst* I digested pBluescript SK+ and screening the resulting size-selected library with the c61-1 insert. One clone obtained from this screening, P6.4 contains a 6.4 kb *Pst* I insert that hybridizes to the c61-1 insert.

A restriction map for P6.4 was constructed using the restriction endonucleases *Sac* I, *Pst* I, and *Sal* I (figure 9). Southern blots performed on *Pst* I-*Sal* I digested P6.4 DNA and probed with the c61-1 insert demonstrated that all of the H2B encoding sequence is located within the 3.5 kilobase *Pst* I-*Sal* I

fragment (data not shown). This *Pst* I-*Sal* I fragment was sub-cloned into pBluescript SK+ for sequence analysis, and this subclone was designated PS3.5. Deletions were generated in both directions, using the enzyme exonuclease III, in order to sequence the entire fragment. Sean Stack in our laboratory assisted in this portion of the sequencing. The final sequence contained 3,524 bp, and encoded three tandemly arranged repeats of the H2B gene (figure 10), designated copies A, B and C. The first two repeats contain a region identical in sequence to cDNA c18, and the third repeat is identical to cDNA c43. There are only three base changes when the intergenic region between copies A and B is compared to the intergenic region between copies B and C. There is an additional T at position 1143 (underlined in figure 10), and there is a base change between intergenic regions A-B and B-C: G at 1161 and A at 2077 (marked by asterisks in figure 10).

Densitometry

From earlier Southern blots, it was determined that there are multiple copies of this gene. A *Pst* I digest of genomic DNA yields three bands (A, B, and C, figure 3) which hybridize to the cDNA c18. P6.4 is a cloned copy of band b. A Southern blot of *Pst* I digested genomic DNA was used for densitometry to estimate the number of H2B genes within the genome. The relative integrated areas under peaks A, B and C are 1.2, 3.0, and 7.4, respectively (figure 11), normalizing peak B to 3.0 to account for the three sequenced H2B genes. These results indicate that these fragments contain 1, 3 and approximately 7 respectively copies of the H2B gene with a total of 11 or 12.

Chromosomal Analysis

The next objective was to determine the arrangement of the clusters on the chromosomes. For this purpose, individual chromosomes of *L. enriettii* were resolved by pulsed field gel electrophoresis (42). The gels were stained

with ethidium bromide and photographed (figure 12) and were then transferred to filters for hybridization with radiolabelled probe. The results indicate that the H2B probe hybridized to two different chromosomes. From the relative intensity of the two hybridization signals, it appears that the higher molecular weight chromosome contains the 7 copy cluster and the 3 copy cluster, with the single copy being on a chromosome of lower molecular weight.

Inhibition of DNA Synthesis and Histone H2B mRNA Levels

In other eukaryotes, cell-cycle regulated histone mRNA levels are greatly affected by inhibition of DNA synthesis. In many higher eukaryotes the addition of DNA synthesis inhibitors to logarithmically growing cells, greatly reduces the level of histone cell-cycle dependent messages (20,25), as determined by Northern gels. This is believed to be analogous to the termination of DNA synthesis at the end of S phase. Inhibition of DNA synthesis decreases the rate of histone gene transcription and also decreases the half-life of the histone mRNA (60,61)

The most common DNA synthesis inhibitors used in these experiments are hydroxyurea and aphidicolin. Hydroxyurea is known to block ribonucleotide reductase in both eukaryotic and prokaryotic organisms (50), which leads to a block in DNA synthesis by inhibiting synthesis of deoxynucleotide precursors. Aphidicolin is known to specifically block DNA polymerase α and δ in eukaryotic organisms (48,49,71). It has recently been shown that DNA synthesis in two species of *Leishmania* (*L. tropica and L. donovani*) is sensitive to both hydroxyurea and aphidicolin. This was determined by monitoring DNA synthesis, in the presence and absence of

inhibitors, by following the incorporation of ^{3}H-hypoxanthine into DNA and RNA (44).

Before looking at the response of histone H2B levels to these drugs, I demonstrated that hydroxyurea and aphidicolin do inhibit DNA synthesis in *L. enriettii*. In the absence of drugs, DNA synthesis was linear for at least 8 hours, following addition of labelled DNA precursor (figure 13A). However, 10 μM aphidicolin inhibited DNA synthesis by 89% and 5 mM hydroxyurea inhibited DNA synthesis by 95% at 10 hours post addition of either drug (figure 13B). These concentrations of drug are similar to those required to inhibit DNA synthesis in higher eukaryotes and in *L. tropica* and *L. donovani*. It is possible that the remaining DNA synthesis may be due to kinetoplast DNA polymerase, as it is known that mammalian mitochondrial DNA polymerase γ is insensitive to aphidicolin (51).

To test the effect of DNA synthesis inhibition on histone H2B mRNA levels, total RNA was isolated at time 0 and then at two hour intervals, beginning at t=8 hr, as described, and analyzed on Northern blots (figure 14). Filters were hybridized with radiolabelled c18 probe to identify H2B mRNA, and also with two control probes, Pro-1 and β-tubulin. Levels of Pro-1 and β-tubulin mRNA should not be affected by DNA synthesis and should therefore remain constant. However, the results for histone H2B message levels were very different from those obtained with higher eukaryotes. In most eukaryotic organisms, the effect of these inhibitors on histone mRNA levels is detectable within minutes: when 5 mM hydroxyurea is added to mouse myeloma cells, levels of histones decrease to 50% within 20 minutes of addition, and are approximately 10% of control level within one hour of drug addition (23). In contrast, with *L. enriettii* no decrease in histone H2B mRNA level was observed for 8-12 hours. In the presence of 5 mM hydroxyurea, the histone message

level remained constant, whereas control messages were shown to slightly decrease. In the presence of 10 μM aphidicolin, the histone message level actually increased slightly over time, while the controls remained constant.

S1 Mapping Experiments: Possible Precursor Transcripts

Prior to initiating the sequence analysis described above, it was necessary to determine if the P6.4 clone contained the entire coding region or just the 3' or 5' portion. This was determined by S1 mapping. In this procedure double stranded DNA was hybridized with RNA under conditions where DNA duplexes melt, but RNA-DNA duplexes are stable (approximately 5°C above melting temperature in formamide buffers) (63). To do mapping experiments, it was necessary to first determine the melting temperature of the P6.4 clone. This was done by setting up reactions containing linearized double stranded P6.4 DNA, and then incubating at temperatures ranging from 40-65°C to determine the temperature at which melting occurred. S1 nuclease was added, and single strandedness (ability to be digested by S1 nuclease) was determined by Southern blotting the digested samples and probing with c61-1 (data not shown). It was determined that the melting temperature for this piece of DNA is approximately 57°C (i.e. P6.4 DNA was digested by S1 nuclease when incubated above 57° but not when incubated below 57°), therefore the appropriate temperature to use for the S1 experiments is 62°C.

For the S1 mapping experiments, double stranded, linearized P6.4 DNA was first denatured and then allowed to anneal with total promastigote RNA at 62°C. S1 nuclease was added and single stranded nucleic acid was digested. Phenol/chloroform, ethanol precipitated samples were analyzed on alkaline agarose gels. By running gels under alkaline conditions, any RNA which hybridized to DNA would be degraded, and the protected DNA fragments could then hybridize to appropriate probe. The resulting Northern blot showed a

protected DNA fragment that hybridized to c61-1 that was about 700 bp long, the length of the H2B mRNA. Therefore, P6.4 probably contains all of the coding sequence (figure 15).

A second result obtained from the S1 experiment, is that there are two larger protected bands which hybridize to c61-1 (arrows, figure 14). It is possible that these protected fragments represent precursors, and will be discussed in more detail in the Discussion section.

DISCUSSION

Developmentally Regulated Genes in *Leishmania*

Leishmania spp. have a two host life cycle. When the parasites are in the mammalian host they live intracellularly, are ovoid and lack a flagellum. When ingested by the sandfly, the parasites undergo a morphological transformation and become elongated, have flagella and live extracellularly. The two stages differ in their morphology, biochemistry, and physiological environments. This implies that there are developmentally regulated genes which are involved in the transformations and adaptations of the two stages. Little is currently known about gene expression or gene regulation in *Leishmania*. The long term goal of this work is to identify several regulated genes which can be used to study mechanisms of gene regulation and expression.

One candidate for a developmentally regulated gene was encoded by the c61-1 cDNA clone. The mRNA accumulates approximately 10-fold more in promastigotes compared to amastigotes. The major objective of this project was to study the structure of the gene encoding c61-1 and to identify a protein product, if possible.

Identification of a *Leishmania* Histone H2B

The identity of this developmentally regulated gene was determined by sequencing several cDNAs and a genomic fragment. An open reading frame, coding for 111 amino acids was identified, and a homology search of the NBRF protein database identified 23 histone H2B protein sequences which were homologous to this open reading frame. BESTFIT analysis aligned the deduced sequence with the histone H2B sequence of a sand sea urchin, with an identity of 48%. Histones in general are highly conserved, with more variation occuring in H1, H2A, and H2B than in H3 and H4. This is the most divergent H2B

sequence known, which is consistent with the large phylogenetic distance between *Leishmania* and other species.

A comparison of the *Leishmania* histone with the known histone H2B protein sequences, indicates both similarities and differences. It is similar to other H2Bs identified with respect to size, molecular weight and conservation at the carboxy terminus of the protein. While the average number of amino acids for H2B is 120 (14), this *Leishmania* histone encodes a 111 amino acid protein. From sequence alignment, it was determined that the *Leishmania* protein is truncated at the amino terminus, compared to the sea urchin histone. The molecular weights of previously identified H2Bs have ranged from 13-17 kilodaltons; while the predicted molecular weight of the *Leishmania* H2B is 12.3 kilodaltons. As with H2Bs already identified, the *Leishmania* H2B is most homologous to other H2Bs at the C-terminus. The significance of this is unclear. It is known that the charged amino-terminus is involved in binding DNA, and it is believed that the apolar central region of the protein contacts other histones in the core of the nucleosome (14). It has been suggested that the C-terminus may also contact DNA in some cases (14). Histones are classified in part by their percent lysine and arginine content. H2Bs are approximately 17% Lys and 7% Arg. In contrast, the *Leishmania* H2B is 9% Lys and 10% Arg. This is the most primitive eukaryote (1) from which an H2B has been isolated.

Structure of *Leishmania* Histone H2B mRNAs

Two of the cDNAs isolated, c18 and c43, were close to full length. Both contained the entire protein coding region and the poly-A tail, and several bases of 5' untranslated region. Sequence analysis identified several features similar to both higher eukaryote cell-cycle regulated histones and yeast histones (figure 16).

There are several structural features which higher eukaryote, cell-cycle regulated histones share. They do not contain intervening sequences or poly-A tails, but instead contain a highly conserved stem-loop at the 3' end of the mRNA. The stem-loop is an essential part of the cell-cycle regulation of higher eukaryote histones; it is both necessary and sufficient for destabilization of histone mRNAs outside of the S phase. Analysis of the 3' termini of the two cloned cDNAs has identified a region in each mRNA with the potential to form a stem-loop. However, the stem-loop in the c18 cDNA is quite different in sequence from the stem-loop in the c43 cDNA. In both cDNAs, the stem consists of 7 nucleotides with a single mismatch, and there is a four basepair loop. Neither sequence is conserved with respect to the other histone stem-loops identified so far. The free energy of the c18 stem-loop, which can form a wobble base pair at the G-U mismatch, was calculated using the Stemloop algorithm as -6.4 kcal/mole, and the c43 stem-loop, which has a G-G mismatch, as 0 kcal/mole. For comparison, the free energy of the higher eukaryote histone stem-loop was calculated to be -7.9 kcal/mole. Therefore, the c18 stem-loop should be capable of forming *in vivo*. The c43 stem-loop may exist in an equilibrium between a linear state and a folded form and could be held in one state by a protein or other macromolecule. It has been demonstrated for sea urchins that the stem-loop structure is recognized by some factor which is involved in processing (to be discussed later). The role of these putative structures *in vivo* in *Leishmania* has not been studied, but their regulatory role in other organisms and their proximity to the poly-A tail suggests that they may have a role in either stability of the message or as a signal for 3' end formation and/or polyadenylation of histone mRNAs or both. Polyadenylation signals for other *Leishmania* mRNAs have not yet been identified.

Yeast histone mRNAs are polyadenylated and lack any sort of stem-loop structure at the 3' terminus, yet as already discussed in the Introduction, they exhibit a dependence on DNA synthesis. Furthermore, basal histone mRNAs of higher eukaryotes are also polyadenylated and lack a stem-loop, but they are not cell-cycle regulated. Both sequence analysis of the *Leishmania* cDNAs and Northern blots of *Leishmania* poly A+ RNA have demonstrated that these histone mRNAs are also polyadenylated. Hence, *Leishmania* H2B mRNAs are unique in containing both poly-A tails and stem-loops in the same mRNA.

Genomic Arrangement of Histone H2B Genes

The genomic arrangement of *Leishmania* H2B genes is unusual with respect to histones, which are usually arranged in repeating clusters of all five histone genes, rather than as tandem repeats of a single type of histone gene. However, many genes in *Leishmania* or the related trypanosomes are arranged in homogeneous tandem repeats. The PS3.5 genomic fragment which was analyzed, contained three tandemly arrayed copies of H2B, two of which were identical to c18, and the third copy was identical to c43. The protein coding regions of these three genes are identical, but the 3' untranslated regions of c18 and c43 diverge from each other (figure 4). The intergenic regions are identical except for one insertion, and a substitution of a G for an A.

Densitometry analysis of Southern blots, leads to the conclusion that there are approximately 11 H2B copies, which are arranged in three clusters. Based on sequence analysis of three different histone cDNAs, it is clear that at least three of the 11 copies are expressed and that the protein coding regions are highly conserved.

Analysis of *Leishmania* chromosomes using pulsed field gel electrophoresis, has identified at least 23 chromosomes (64). PFGE gels showed

that the three histone H2B clusters are localized on two chromosomes, with the 7 copy- and the 3 copy-clusters on the same chromosome, based on intensity of hybridization, and the single copy on the lower molecular weight chromosome. The possibility that the two multi-copy clusters are linked and are merely separated by a *Pst* I site cannot be ruled out.

Regulation of Histone H2B Expression

The question of why histone H2Bs are developmentally regulated remains to be answered. Developmental regulation of histones has been observed in other systems. An example of this is sea urchin. The predominant histones of the egg and early embryo are the α-variants. There is a second set, cleavage-stage histones, which are expressed for only a few hours immediately following fertilization. Later, embryos express yet a third set of histones, γ-variants (65). The significance of differential expression of histones during embryogenesis is not clear.

Clearly the bulk of H2B is differentially expressed between the two life cycle stages of *Leishmania*. The 3 genes investigated here encode the same protein. We do not know for certain whether other genes might encode alternate histones which are differentially expressed.

There is some evidence that histones can be involved in transcription regulation, by being ubiquitinated, but it is at best uncertain (69,62). The addition of ubiquitin in general targets proteins for degradation. In the case of histones, it can be argued that by targeting a histone for destruction, you are actually opening up a region of DNA for transcription. The histone which has been seen most often ubiquitinated is in fact H2B (69,62). There is no evidence that this is occurring in *Leishmania* with the H2B in question; however, it might provide one explanation for the differential expression between the two parasite stages. It may be that not all of the H2Bs that are detected are

regulated, but because of the conservation, it is impossible to single out the one(s) which is developmentally regulated.

Another possible explanation for the difference in abundance of H2B mRNA in promastigotes compared to amastigotes could be that it is not developmental regulation *per se*, but a reflection of different doubling times in the two life cycle forms. Since the promastigotes are grown in tissue culture under optimal conditions, compared to amastigotes grown in Guinea pig lesions, it is reasonable to assume that promastigotes are doubling at a faster rate, and hence have a greater demand for newly synthesized histone protein. It has been determined that the doubling time of promastigotes in tissue culture is 6 hours. Although the doubling time of amastigotes in Guinea pig lesions cannot be determined, a division time of 36 hr has been measured (58) for *L. mexicana* amastigotes growing inside tissue culture macrophages. So, the greater abundance of H2B message seen in the promastigote stage could simply be the result of more rapid division of promastigotes.

In other systems, the most common approach to determining histone mRNA dependence on DNA synthesis is to monitor histone message levels in the presence of DNA synthesis inhibitors. In the case of cell-cycle regulated histones of higher eukaryotes, mRNA levels dramatically decrease within minutes of addition of a DNA synthesis inhibitor. Similar experiments were performed with *L. enriettii*, using the DNA synthesis inhibitors hydroxyurea and aphidicolin. I first demonstrated that DNA synthesis in *L. enriettii* is strongly inhibited by these drugs, as was shown previously for *L. tropica* and *L. donovani*. I then treated *L. enriettii* with these two drugs for periods ranging from 8 to 12 hours, isolated RNA and performed Northern blots to determine the effect on histone H2B mRNA levels. Contrary to expectation, the histone H2B mRNA levels are not affected as seen in other eukaryotic

organisms. The H2B mRNA levels stayed constant or actually appeared to increase slightly, as in the case of aphidicolin. Similar experiments monitoring H2B mRNA levels from 0 to 4 hours after addition of inhibitors also failed to show any decrease in histone message (data not shown). There are several possible explanations for this result: (1) This is not a cell-cycle regulated histone, but in fact, is a basal, or variant histone, which is constitutively expressed independently of DNA synthesis or cell cycle phase (2) The factor(s) necessary for regulating H2B message levels is not present in S phase. When inhibitors of DNA synthesis are added, DNA synthesis is blocked and presumably the cells are arrested in S phase. In the absence of regulatory factors, the H2B mRNAs might remain constant or actually begin to accumulate. (3) These H2B mRNAs are regulated during cell cycle, but they respond to some signal other than DNA replication. In yeast, the half-life of histone H2B mRNA is regulated during the cell-cycle, but the half-life is not affected by inhibition of DNA synthesis. Apparently yeast employ some signal other than DNA replication to regulate histone H2B mRNA half-life. A similar mechanism may operate in *Leishmania*.

One *caveat* of any drug experiment is that the drug may induce alterations in cellular processes other than those under study. These additional alterations could result in artefacts that interfere with the process of interest. However, aphidicolin is a very specific inhibitor of DNA polymerases α and δ and does not affect transcription or translation (71), suggesting that the failure of H2B mRNA levels to decrease in the presence of aphidicolin is not due to a drug artefact.

Possible Polycistronic Transcripts of Histone H2B Genes

Several lines of evidence (11,12) suggest that tandemly arrayed genes in trypanosomes are being transcribed as a single transcript and then processed

down into individual transcripts (UV inactivation, heat shock). With the systems used previously it was necessary to block processing in some way to see putative precursor messages. When doing S1 mapping of the P6.4 genomic fragment, which contains three tandemly repeated copies of the H2B gene, several higher molecular weight bands were present in the lanes which contained DNA, RNA and S1 (figure 15). One possible explanation for these bands is that they represent polycistronic precursors containing two or three units of the mRNA. Bands of greater than unit size are also present on *Leishmania* total RNA Northerns probed with c18 or c43 (figure 2, lanes 3 and 4, arrowheads), and these could also be precursors of mature mRNA. However, more detailed analysis would be necessary to determine whether these higher molecular weight bands are actual precursors.

Homology to Eukaryotic 3' Processing Signal

A third type of histone mRNA regulation seen in higher eukaryotes involves processing to form the mature 3' end. A sequence involved in the processing has been identified and is located 3' to the stem-loop in mammalian histone pre-mRNAs. In sea urchins, this sequence has been shown to base pair with the U7 snRNA and is essential for mRNA processing; base changes or deletions of this sequence fail to give rise to mature message (66). An additional factor makes contact with the hairpin loop, and is also essential for 3' formation (67). I have identified a sequence downstream from the third copy in the genomic cluster PS3.5 (fig. 10), which is similar to the consensus sequence for processsing:

Sea urchin CAAGAAAGA

Leishmania GAAGAAATT

It is possible that this sequence is involved in the initial processing of putative precursors. However, to define a possible functional role, it would be

necessary to delete or mutagenize this sequence and transform the altered construct back into *Leishmania* parasites.

Future Directions

Several features of *Leishmania* histone H2B have been described here, which make it unique from previously identified H2Bs. However, until recently it was not possible to test these structures or sequences for function. Stable transformation of *L.enriettii* is now available (68). It will be interesting to fuse targeted sequences to non-histone genes and perform mutations or deletions on the putative stem-loops or 3' consensus sequences. It would then be possible to test for proper 3' end formation by S1 or ribonuclease mapping. It will still be difficult to study signals involved in cell-cycle regulation in *Leishmania* because: (1) there is no known way to synchronize *Leishmania* cells, and (2) H2B mRNA is not reponsive to DNA synthesis inhibitors, so this cannot be used as a substitute for cell-cycle regulation.

Summary

This project began with a differentially expressed gene of unknown function. I have since sequenced both cDNAs and genomic clones encoding this gene. DNA sequence analysis has allowed the identification of this gene as a histone H2B. mRNA sequence shows that it is unique as compared to previously identified histone H2Bs. Unlike any other known H2B, it contains both a poly-A tail and a novel stem-loop. Genomic sequence analysis shows that the genes are arranged in tandem repeats of similar but not identical sequence. From Southerns and pulsed field gels, I have shown that there are three clusters on two different chromosomes. The regulation of *Leishmania* histone H2B genes is different from the cell-cycle regulated histone genes of other eukaryotes; specifically, levels of histone H2B mRNA are not decreased by inhibitors of DNA synthesis . And finally, I have shown preliminary

evidence consistent with tandemly arrayed H2B genes being transcribed as polycistronic transcripts.

REFERENCES

1. Sogin, M. L. et al., (1986) Evolutionary diversity of small-subunit rRNA genes. PNAS 83, 1383-87.
2. Zuckerman, A. and Lainson, R. (1979) In: Parasitic Protozoa (Kreier, J.P., ed.), Vol. 1, 57-133, Academic Press, New York.
3. Glew, R. et al., (1988) Biochemistry of the *Leishmania* species. Microbio. Rev., 52, 412-32.
4. Benne, R. et al., (1986) Major transcript of the frameshifted *cox II* gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. Cell 46, 819-26.
5. Feagin, J. E., et al., (1987) Developmentally regulated addition of nucleotides within the apocytochrome b transcripts in *Trypanosoma brucei*. Cell 49, 337-45.
6. Murphy, W. J. et al., (1986) Identification of a novel Y branch structure in Trypanosome mRNA processing: evidence for trans splicing. Cell 47, 517-25.
7. Sutton, R. E. and J. C. Boothroyd (1986) Evidence for trans splicing in trypanosomes. Cell 47, 527-35.
8. Landfear, S. M., et al., (1983) Tandem arrangement of tubulin genes in the protozoan parasite *Leishmania enriettii*. Mol. Cell. Biol. 3, 1070-76.
9. Cairns, B. R., et al., (1989) Developmentally regulated gene from *Leishmania* encodes a putative membrane transport protein. PNAS 86, 7682-86.
10. Yost, H. J., and S. Lindquist (1986) RNA splicing is interrupted by heat shock and is rescued by heat shock protein synthesis. Cell 45, 185-93.

11. Muhich, M. L. and J. C. Boothroyd (1988) Polycistronic transcripts in trypanosomes and their accumulation during heat shock: evidence for a precursor role in mRNA synthesis. Mol. Cell. Biol., 8, 3837-46.

12. Johnson, P. J., et al., (1987) Inactivation of transcription by uv irradiation of *T. brucei* provides evidence for a multicistronic transcription unit including a VSG gene. Cell 51, 273-81.

13. Maxson, R., et al., (1983) Expression and organization of histone genes. Ann. Rev. Genet. 17, 239-77.

14. Lewin, B. (1980) Gene Expression, Vol.2, John Wiley & Sons, New York.

15. Hentschel, C. C. and M. L. Birnstiel (1981) The organization and expression of histone gene families. Cell 25, 301-13.

16. Bravo, R., and J. E. Celis (1980) A search for differential polypeptide synthesis throughout the cell cycle of HeLa cells. J. Cell Biol. 84, 795-802.

17. Marashi, F, et al., (1982) Histone proteins in HeLa S3 cells are synthesized in a cell-cycle stage specific manner. Science 215, 683-85.

18. Stein, G. S., et al., (1975) Regulation of cell cycle stage-specific transcription of histone genes from chromatin by non-histone chromosomal proteins. Nature 257, 764-67.

19. Stephens, R. E., et al., (1977) Studies on human histone messenger RNA. J. Biol. Chem. 252, 166-72.

20. Heintz, N., et al., (1983) Regulation of human histone gene expresson: kinetics of accumulation and changes in the rate of synthesis and half life of individual histone mRNAs during the HeLa cell cycle. Mol. Cell. Biol., 3, 539-50.

21. DeLisle, A. J., et al., (1983) Regulation of histone messenger RNA production and stability in serum stimulated mouse 3T6 fibroblasts. Mol. Cell. Biol., 3, 1920-29.

22. Artishevsky, A., et al., (1985) Isolation of a mammalian sequence capable of conferring cell cycle regulation to a heterologous gene. Science 230, 1061-62.

23. Sittman, D. B., et al., (1983) Histone mRNA concentrations are regulated at the level of transcription and mRNA degradation. PNAS 80, 1849-53.

24. Pandey, N. B. and W. F. Marzluff (1987) The stem-loop structure at the 3' end of histone mRNA is necessary and sufficient for regulation of histone mRNA stability. Mol. Cell. Biol., 7, 4557-4559.

25. Gallwitz, D. (1975) Kinetics of inactivation of histone mRNA in the cytoplasm after inhibition of DNA replication in synchronized HeLa cells. Nature 257, 247-48.

26. Graves, R. A., et al., (1987) Translation is required for regulation of histone mRNA degradation. Cell 48, 615-26.

27. Ross, J., (1986) Histone mRNA degradation *in vivo*: the first detectable step occurs at or near the 3' terminus. Mol. Cell. Biol., 6, 4362-71.

28. Luscher, B., and D. Schumperli. (1987) RNA 3' processing regulates histone mRNA levels in a mammalian cell cycle mutant. A processing factor becomes limiting in G1-arrested cells. EMBO J. 6, 1721-26.

29. Lycan, D. E., et al., (1987) Role of transcriptional and posttranscriptional regulation in expression of histone genes in *Saccharomyces cerevisiae*. Mol. Cell. Biol., 7, 614-621.

30. Iovannisci, D. M. and B. Ullman (1983) High efficiency plating method for *Leishmania* promastigotes in semidefined or completely-defined media. J. Parsitol. 69, 633-36.

31. Landfear, S. M., et al., (1986) Transcriptional mapping of *Leishmania enriettii* tubulin mRNAs. Mol. Biochem. Parasit. 21, 235-245.

32. Landfear, S. M. and D. F. Wirth (1984) Control of tubulin expression in the parasitic protozoan *Leishmania enriettii*. Nature 309, 716-17.

33. Landfear, S. M. and D. F. Wirth (1985) Structure of mRNA encoded by tubulin genes in *Leishmania enriettii*. Mol. Biochem. Parasitol., 15, 61-82.

34. Feinberg, A. P. and B. Vogelstein (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132, 6-13.

35. Sanger, F., et al., (1977) DNA sequencing with chain terminating inhibitors. PNAS 74, 5463-67.

36. Henikoff, S. (1984) Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28, 351-59.

37. Maxam, A. M. and W. Gilbert (1980) Sequencing end-labelled DNA with base-specific chemical cleavages. Methods Enzymol. 65, 499-560.

38. Chen, E. Y. and P. H. Seeburg (1985) Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA 4, 165-70.

39. Kraft, R., et al., (1988) Using mini-prep plasmid DNA for sequencing double-stranded templates with Sequenase. Bio-Techniques 6, 544-546.

40. Devereux, J., et al., (1984) A comprehensive set of sequence analysis programs for VAX. Nuc. Acids Res. 12, 387-95.

41. Maniatis, T., et al., (1982) Molecular Cloning, a Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

42. Anand, R. (1986) Pulsed field gel electrophoresis: a technique for fractionating large DNA molecules. Trends in Genet. 2, 278-83.

43. Carle, G. F. and M. V. Olson (1985) An electrophoretic karyotype for yeast PNAS 82, 3756-60.

44. North, T. W. and D. J. Wyler (1987) DNA synthesis in promastigotes of *Leishmania major* and *L. donovani*. Mol. Biochem. Parasitol. 22, 215-21.

45. Gubler, U. and B. J. Hoffman (1983) A simple and very efficient method for generating cDNA libraries. Gene 25, 263-69.

46. Smith, T. F. and M. S. Waterman (1981) Comparison of bio-sequences. Adv. in Appl. Math. 2, 482-89.

47. Pearson, W. R. and D. J. Lipman (1988) Improved tools for biological sequence analysis. PNAS 85, 2444-48.

48. Bucknall, R. A. et al., (1973) Antiviral effects of aphidicolin, a new antibiotic produced by *Cephalosporium aphidicola*. Antimicrob. Agents Chemother. 4, 294-98.

49. Ikegami, S., et al., (1978) Aphidicolin prevents mitotic cell division by interfering with the activity of DNA polymerase-α. Nature 275, 458-60.

50. Wright, J. A. and J. G. Cory (1983) Alternations in the components of ribonucleotide reductase in hydroxyurea-resistant hamster cells. Biosci. Rep. 3, 741-48.

51. Krokan, H., et al., (1979) Involvement of eukaryotic deoxyribonucleic acid polymerases α and γ in the replication of cellular and viral deoxyribonucleic acid. Biochemistry 18, 4431-43.

52. Blum, B., et al., (1990) A model for RNA editing in kinetoplastid mitochondria: "guide" RNA molecules transcribed from maxicircle DNA provide the edited information. Cell 60, 189-98.

53. Busslinger, M., et al., (1980) Ubiquitous and gene specific regulatory 5' sequences in a sea urchin histone gene clone coding for histone protein variants. Nucl. Acids Res. 8, 957-77.

54. Wells, D. E. (1986) Compilation analysis of histones and histone genes. Nucl. Acids Res. 14, r119-49.

55. Huynh, T. V., et al., (1985) in DNA Cloning Techniques: A Practical Approach, ed. Glover, D. (IRL), Oxford), Vol.1, pp. 49-78.

56. Comeau, A. M., et al., (1986) Chromosome location of four genes in *Leishmania*. Mol. Biochem. Parasitol. 21, 161-69.

57. Hereford. L., et al., (1979) Isolation of yeast histone genes H2A and H2B. Cell 18, 1261-71.

58. Chang, K. P. (1980) Human cutaneous *Leishmania* in a mouse macrophage line: Propagation and isoloation of intracellular parasites. Science 209, 1240-42.

59. Perry, K. L., et al., (1987) Trypanosome mRNAs have unusual "cap 4" structures acquired by spliced leader. PNAS 84, 8190-94.

60. Marzluff, W. M. and N. B. Pandey (1988) Multiple regulatory steps control histone mRNA concentrations. Trends in Bioch. Sci. 13, 49-52.

61. Schumperli, D. (1988) Multilevel regulation of replication-dependent histone genes. Trends in Genet. 4, 187-91.

62. Nickel, B. E., et al., (1989) Ubiquitinated histone H2B is preferentially located in transcriptionally active chromatin. Biochem. 28, 959-963,

63. Berk, A. J., and P. A. Sharp (1977) Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease digested hybrids. Cell 12, 721-32.

64. Spithill, T. W. and Samaras, N. (1987) Genomic organization, chromosomal location and transcription of dispersed and repeated tubulin genes in *Leishmania major*. Mol. Biochem. Parasitol. 2, 23-37.

65. Maxson, R., et al., (1983) Expression and organization of histone genes. Ann. Rev. Genet. 17, 239-77.

66. Schaufele, F., et al., (1986) Compensatory mutations suggest that base-pairing with a small nuclear RNA is required to form the 3' end of H3 messenger RNA. Nature 323, 777-81.

67. Vasserot, A. P., et al., (1989) Conserved terminal hairpin sequences of histone mRNA precursors are not involved in duplex formation with the U7 RNA but act as a target site for a distinct processing factor. PNAS 86, 4345-49.

68. Laban, A., et al., (1990) Stable expression of the bacterial *neo*^{r} gene in *Leishmania enriettii*. Nature 343, 572-74.

69. Levinger, L., and A. Varshavsky (1982) Selective arrnagement of ubiquitnated and D1 protein-containing nucleosomes within the *Drosophila* genome. Cell. 28, 375-85.

70. Busslinger, M. (1979) A regulatory sequence near the 3' end of sea urchin histone genes. Nuc. Acids Res. 6, 2997-3008.

71. Spadari, A., et al., (1982) Aphidicolin: a specific inhibitor of nuclear DNA replication in eukaryotes. Trends in Bioch. Sci. 7, 29-32.

Figure 1 Life cycle. *Leishmania* spp. require two hosts to complete their life cycle. The cycle begins with the insect vector, the sandfly, taking a blood meal from a mammal. The promastigotes are released into the skin and internalized by macrophages. The parasites undergo a morphological transformation to the amastigote stage, which lives intracellularly. Amastigotes replicate within the macrophage lysosomes and eventually cause cell rupture. At this point, they can be ingested by surrounding macrophages. This process continues until a sandfly takes a blood meal from an infected host and ingests infected macrophages. The cells are broken down in the gut of the insect, and amastigotes are released which transform back to the promastigote stage. As the parasite matures within the midgut of the sandfly it migrates to the salivary gland of the insect where it can be released during another blood meal.

LIFE CYCLE of—

Leishmania

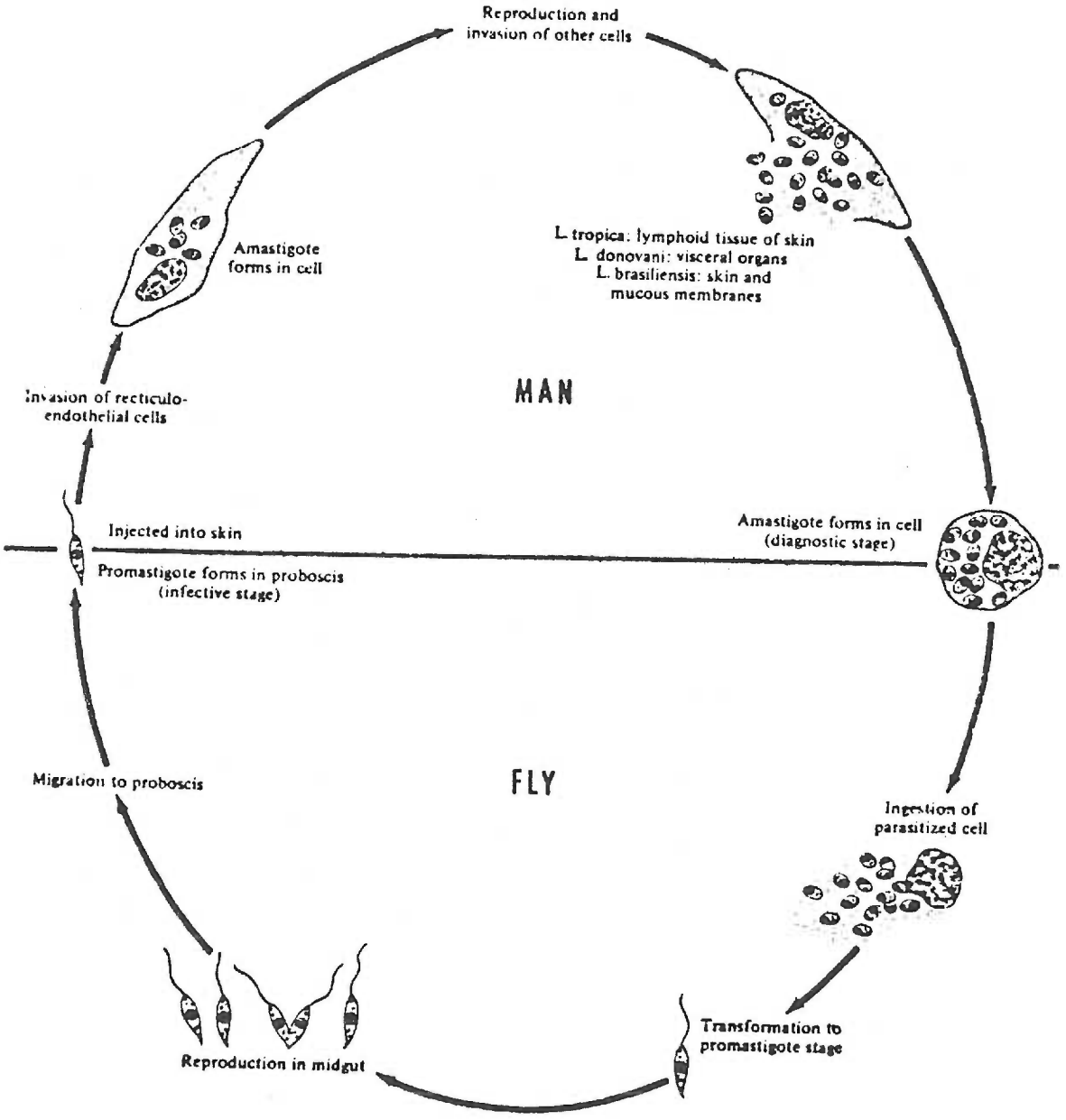


Figure 2 Northern blot of total RNA from promastigote and amastigote stages. Total RNA was isolated from *L. enriettii* amastigotes grown in guinea pig lesions and promastigotes grown in tissue culture. Five μg of total RNA (promastigote lanes 1, 3 and 5, amastigote lanes 2, 4 and 6) were electrophoresed on a 1% agarose-formaldehyde gel, transferred to nitrocellulose and hybridized to c61-1 insert (lanes 1-4). Lanes 1-2 were exposed to film for 1 hour, lanes 3-4, for 18 hours. The filter was rehybridized, without removing H2B probe, with an *L. enriettii* ribosomal RNA probe (56), and re-exposed to film for 15 minutes (lanes 5-6). Arrowheads in lanes 3 and 4 point to bands which are larger than the mature H2B mRNA.

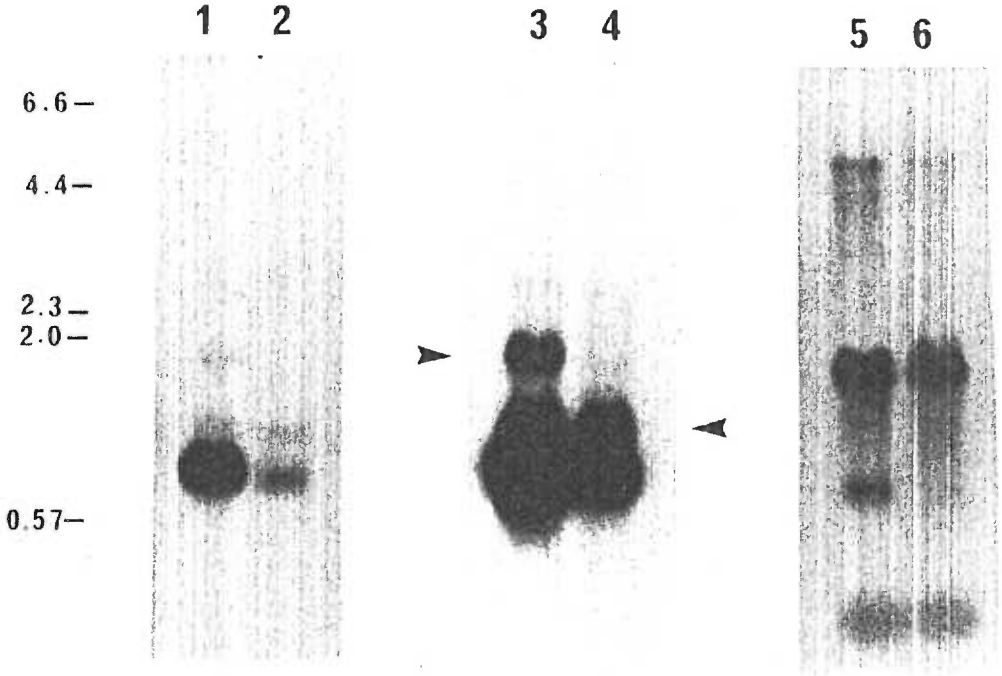

1 2
3 4
5 6
6.6
4.4
2.3
2.0
0.57

Figure 3 Genomic Southern blot hybridized with a histone H2B probe. Approximately 2 μg of genomic DNA from promastigotes was digested with each restriction enzyme (lane 1, *Pst* I; lane 2, *Eco* RI; lane 3, *Hind* III; lane 4 *Bam* HI; lane 5, *Xho* I; lane 6, *Sal* I) and separated on a 1% agarose gel. Following blotting, the filter was hybridized with the radiolabelled insert from the c61-1 cDNA clone. The numbers at the right indicate the sizes and mobilities of fragments *Hind* III digest of λ DNA. Bands a, b and c (arrowheads lane 1) are the 9.4 kilobase, 6.4 kilobase and 3.8 kilobase *Pst* I fragments, respectively, that hybridize to the histone H2B cDNA insert.

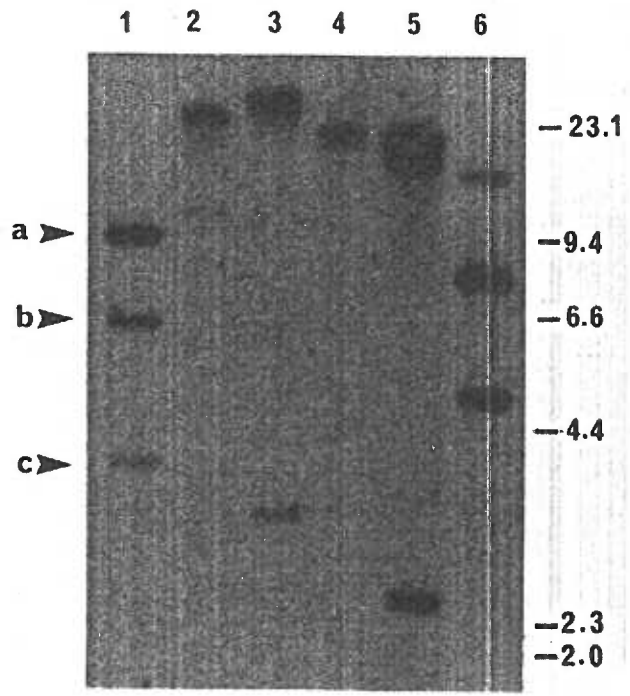

1
2
3
4
5
6
a
b
c
23.1
9.4
6.6
4.4
2.3
2.0

Figure 4 Nucleic acid comparison of cDNA c18 and c43. Two cDNAs were sequenced and are shown here. Gaps in sequence to allow alignment are designated with an asterisk (*). The positions to which oligonucleotide primers A and B and C hybridize are over-lined. Identity between nucleotides is indicated by a vertical bar (|). The deduced amino acid sequence encoded by these cDNAs is shown in figure 5 (protein sequence alone) and in figure 10 (protein sequence aligned with DNA sequence).

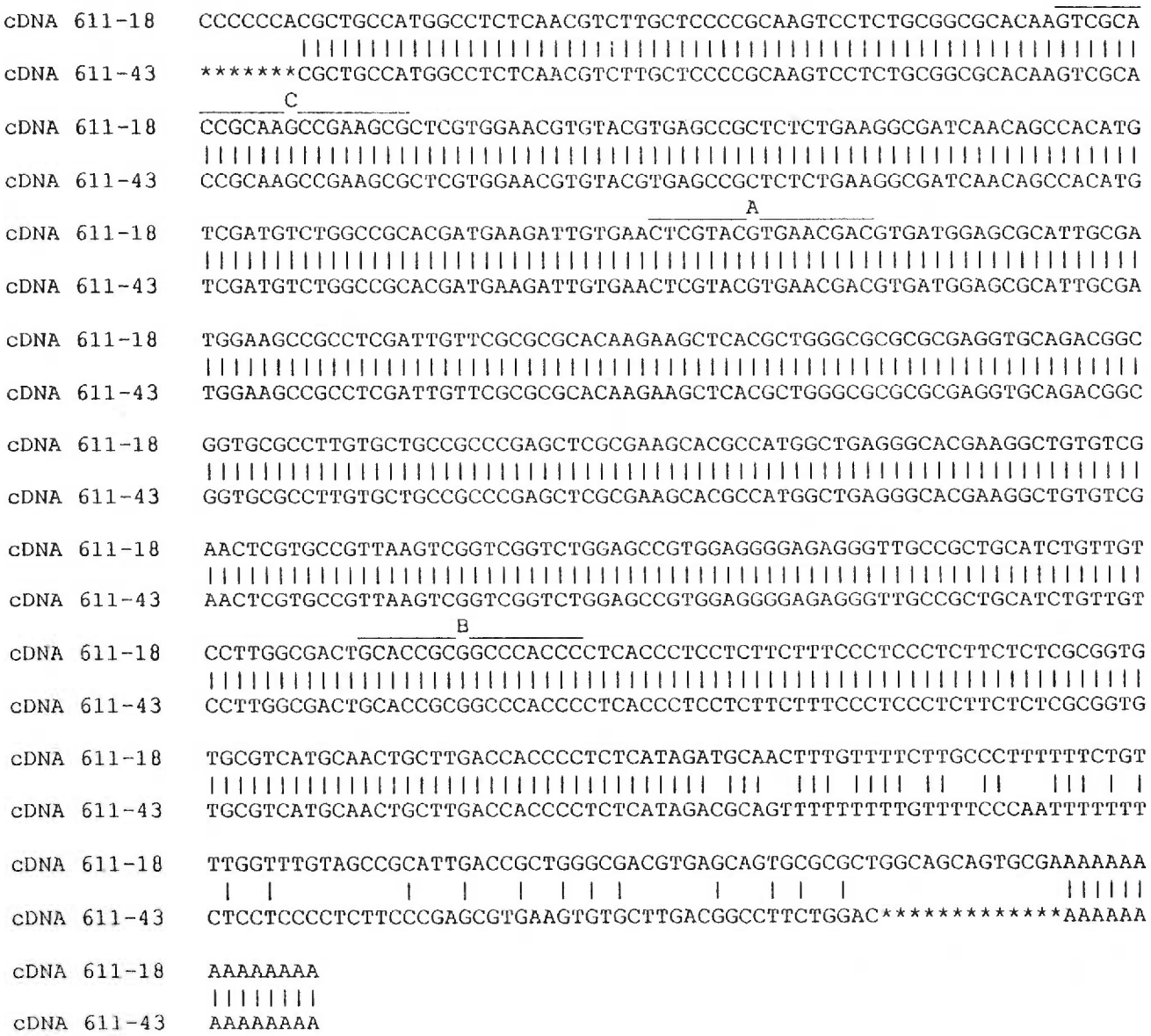

cDNA 611-18 CCCCCCACGCTGCCATGGCCTCTCAACGTCTTGCTCCCCGCAAGTCCTCTGCGGCGCACAAGTCGCA
cDNA 611-43 *******CGCTGCCATGGCCTCTCAACGTCTTGCTCCCCGCAAGTCCTCTGCGGCGCACAAGTCGCA
C
cDNA 611-18 CCGCAAGCCGAAGCGCTCGTGGAACGTGTACGTGAGCCGCTCTCTGAAGGCGATCAACAGCCACATG
cDNA 611-43 CCGCAAGCCGAAGCGCTCGTGGAACGTGTACGTGAGCCGCTCTCTGAAGGCGATCAACAGCCACATG
A
cDNA 611-18 TCGATGTCTGGCCGCACGATGAAGATTGTGAACTCGTACGTGAACGACGTGATGGAGCGCATTGCGA
cDNA 611-43 TCGATGTCTGGCCGCACGATGAAGATTGTGAACTCGTACGTGAACGACGTGATGGAGCGCATTGCGA
cDNA 611-18 TGGAAGCCGCCTCGATTGTTCGCGCGCACAAGAAGCTCACGCTGGGCGCGCGCGAGGTGCAGACGGC
cDNA 611-43 TGGAAGCCGCCTCGATTGTTCGCGCGCACAAGAAGCTCACGCTGGGCGCGCGCGAGGTGCAGACGGC
cDNA 611-18 GGTGCGCCTTGTGCTGCCGCCCGAGCTCGCGAAGCACGCCATGGCTGAGGGCACGAAGGCTGTGTCG
cDNA 611-43 GGTGCGCCTTGTGCTGCCGCCCGAGCTCGCGAAGCACGCCATGGCTGAGGGCACGAAGGCTGTGTCG
cDNA 611-18 AACTCGTGCCGTTAAGTCGGTCGGTCTGGAGCCGTGGAGGGGAGAGGGTTGCCGCTGCATCTGTTGT
cDNA 611-43 AACTCGTGCCGTTAAGTCGGTCGGTCTGGAGCCGTGGAGGGGAGAGGGTTGCCGCTGCATCTGTTGT
B
cDNA 611-18 CCTTGGCGACTGCACCGCGGCCCACCCCTCACCCTCCTCTTCTTTCCCTCCCTCTTCTCTCGCGGTG
cDNA 611-43 CCTTGGCGACTGCACCGCGGCCCACCCCTCACCCTCCTCTTCTTTCCCTCCCTCTTCTCTCGCGGTG
cDNA 611-18 TGCGTCATGCAACTGCTTGACCACCCCTCTCATAGATGCAACTTTGTTTTCTTGCCCTTTTTTCTGT
cDNA 611-43 TGCGTCATGCAACTGCTTGACCACCCCTCTCATAGACGCAGTTTTTTTTTGTTTTCCCAATTTTTTT
cDNA 611-18 TTGGTTTGTAGCCGCATTGACCGCTGGGCGACGTGAGCAGTGCGCGCTGGCAGCAGTGCGAAAAAAA
cDNA 611-43 CTCCTCCCCTCTTCCCGAGCGTGAAGTGTGCTTGACGGCCTTCTGGAC**********AAAAAA
cDNA 611-18 AAAAAAAA
cDNA 611-43 AAAAAAAA

Figure 5 Maximal alignment of histone H2B from the sea urchin *Psammechinus miliaris* (top) and *Leishmania enriettii* (bottom) using the BESTFIT algorithm. Vertical bars (|) indicate amino acids which are identical between the two proteins. A single gap is introduced, marked with an asterisk (*), to allow maximal alignment. The numbers at the right mark the amino acid position of each sequence, starting with the amino terminal methionine.

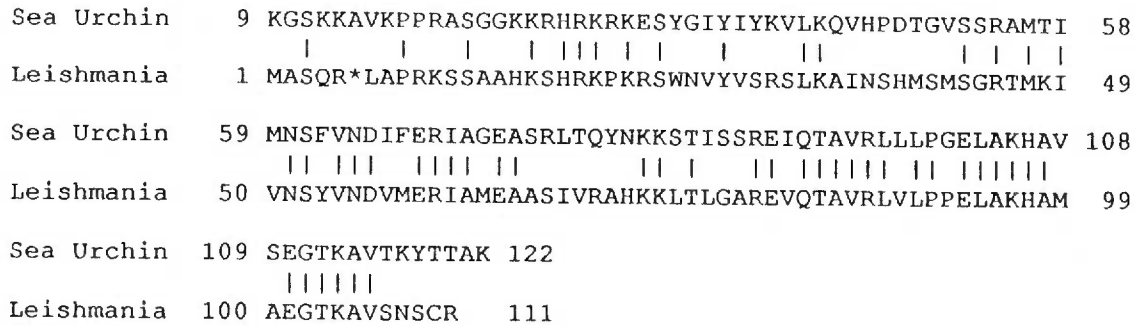
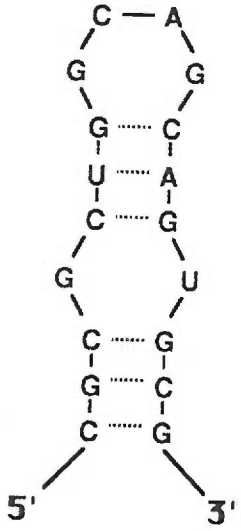

Sea Urchin 9 KGSKKAVKPPRASGGKKRHRKRKESYGIYIYKVLKQVHPDTGVSSRAMTI 58
Leishmania 1 MASQR*LAPRKSSAAHKSHRKPKRSWNVYVSRSLKAINSHMSMSGRTMKI 49
Sea Urchin 59 MNSFVNDIFERIAGEASRLTQYNKKSTISSREIQTAVRLLLPGELAKHAV 108
Leishmania 50 VNSYVNDVMERIAMEAASIVRAHKKLTLGAREVQTAVRLVLPPELAKHAM 99
Sea Urchin 109 SEGTKAVTKYTTAK 122
Leishmania 100 AEGTKAVSNSCR 111

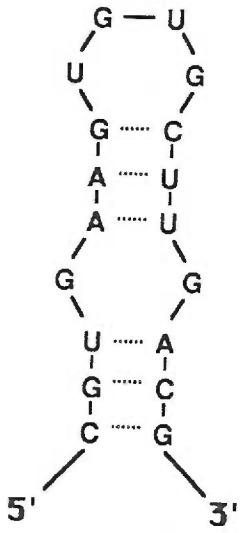
Figure 6 Primer extension analysis. Primer extension products generated when 0.4 pmoles of end-labelled primer C (figure 4) were hybridized to 3 µg of promastigote polyadenylated RNA and extended using reverse transcriptase were analyzed on a 6% acrylamide/8 M urea sequencing gel (lane +). Lane P is end-labelled primer alone, and lane - is a control sample treated the same as (+) but to which no RNA was added. Lanes G, A, T, and C correspond to the dideoxynucleotide sequence reactions of a *Sac* I fragment from genomic clone PS3.5 (figure 9) also primed with oligonucleotide C. The thick arrow designates the primary extension product, and the thin arrow marks the start of translation.

G A T C P

Figure 7 Comparison of stem-loop structures located at the 3' termini of mRNA. Dotted lines designate Watson-Crick base pairs within the stem. Both structures contain a single mis-match in the middle of the stem, indicated by the bulge, and a four base-pair loop. However, the G-U pair in c18 can form a wobble base pair. The poly-A tail begins immediately after the c18 stem-loop and 13 bases down from the c43 stem-loop.



c18



c43

Figure 8 Presence of histone H2B mRNA in the polyadenylated RNA fraction. Five µg of total RNA (lanes T), 2 µg of polyadenylated RNA (lanes A+) and 5 µg of non-polyadenylated RNA (lanes A-) were separated on 1% agarose-formaldehyde gels and transferred to nylon filters. The filter on the left was hybridized with a mixture of radiolabelled probes from the Pro-1 transporter gene and from the histone H2B gene (c43 cDNA clone). The filter on the right was hybridized with a cloned ribosomal RNA probe (56) from *L. enriettii*. The positions of the Pro-1 transporter mRNA (3.1 kilobases) and the H2B mRNA (750 nucleotides) are indicated by arrowheads. Courtesy of David Stein.

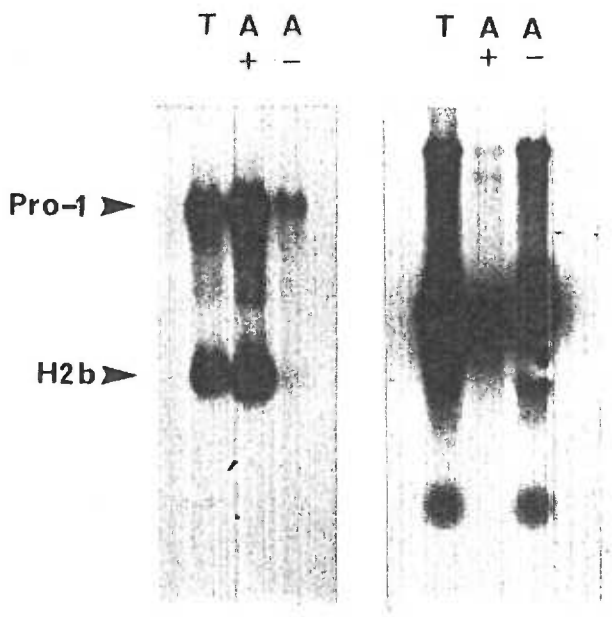

T A A
+ −
T A A
+ −
Pro-1
H2b

Figure 9 A restriction map of the H2B genomic clone P6.4. The 6.4 kb genomic *Pst* I fragment (band b, figure 3) is shown on top. P = *Pst* I, S = *Sal* I, and Sac = *Sac* I. PS3.5 is the 3.5 kb *Pst* I-*Sal* I fragment which was used for sequencing. The solid bars represent the histone H2B encoding regions; A, B, and C designate the first, second and third copies of the gene. The cDNA c18 is identical to copies A and B, and c43 is identical to copy C.

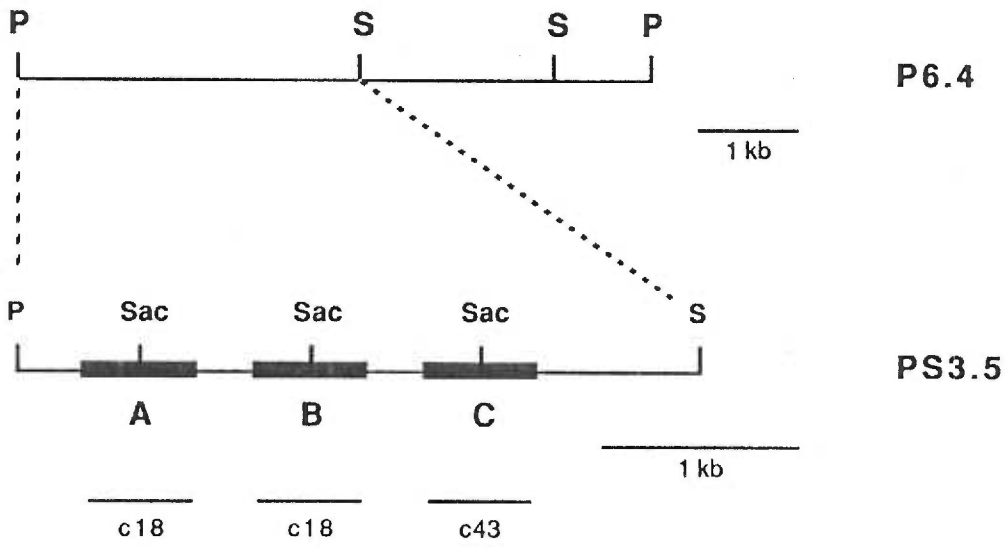
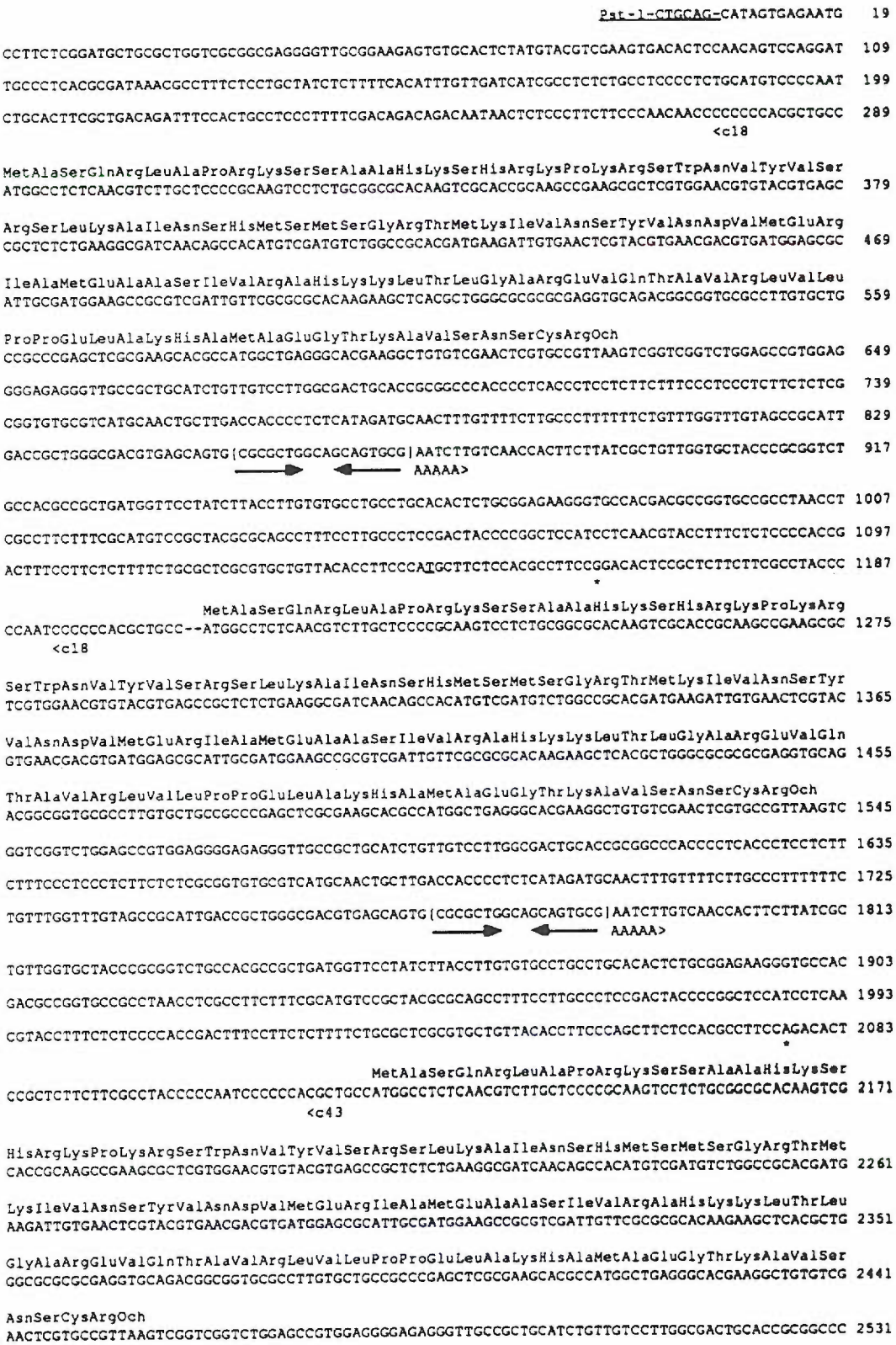

P
S
S
P
P6.4
1 kb
P
Sac
Sac
Sac
S
PS3.5
A
B
C
1 kb
c18
c18
c43

Figure 10 Sequence of genomic clone PS3.5. The genomic sequence is indicated from base 1 to 3524 (includes the *Pst* I and *Sal* I cloning sites as shown). The encoded amino acids are displayed over each triplet in the three coding regions, including the ochre (Och) termination codons. The 5' positions of cDNAs are indicated by <c18 and <c43 directly below the first base of the cDNA. The 3' ends of the cDNAs are indicated by the AAAAA> symbol, representing the position of the start of the poly-A tail. The stem-loop structures are enclosed in brackets, and the stems are underlined with opposing arrows. The T insertion at postion 1143 is underlined, and the asterisks beneath the G at position 1161 and the A at position 2077 mark the G to A transition at these positions in the intergenic region. The sequence homologous to the processing signal is marked with a heavy line (see Discussion). The dashes (-) proceeding start codons within the genomic sequence are introduced so that codons are not divided between lines.


Pst-1-CTGCAG-CATAGTGAGAATG 19
CCTTCTCGGATGCTGCGCTGGTCGCGGCGAGGGGTTGCGGAAGAGTGTGCACTCTATGTACGTCGAAGTGACACTCCAACAGTCCAGGAT 109
TGCCCTCACGCGATAAACGCCTTTCTCCTGCTATCTCTTTTCACATTTGTTGATCATCGCCTCTCTGCCTCCCCTCTGCATGTCCCCAAT 199
CTGCACTTCGCTGACAGATTTCCACTGCCTCCCTTTTCGACAGACAGACAATAACTCTCCCTTCTTCCCAACAACCCCCCCCACGCTGCC 289
<c18
MetAlaSerGlnArgLeuAlaProArgLysSerSerAlaAlaHisLysSerHisArgLysProLysArgSerTrpAsnValTyrValSer
ATGGCCTCTCAACGTCTTGCTCCCCGCAAGTCCTCTGCGGCGCACAAGTCGCACCGCAAGCCGAAGCGCTCGTGGAACGTGTACGTGAGC 379
ArgSerLeuLysAlaIleAsnSerHisMetSerMetSerGlyArgThrMetLysIleValAsnSerTyrValAsnAspValMetGluArg
CGCTCTCTGAAGGCGATCAACAGCCACATGTCGATGTCTGGCCGCACGATGAAGATTGTGAACTCGTACGTGAACGACGTGATGGAGCGC 469
IleAlaMetGluAlaAlaSerIleValArgAlaHisLysLysLeuThrLeuGlyAlaArgGluValGlnThrAlaValArgLeuValLeu
ATTGCGATGGAAGCCGCGTCGATTGTTCGCGCGCACAAGAAGCTCACGCTGGGCGCGCGCGAGGTGCAGACGGCGGTGCGCCTTGTGCTG 559
ProProGluLeuAlaLysHisAlaMetAlaGluGlyThrLysAlaValSerAsnSerCysArgOch
CCGCCCGAGCTCGCGAAGCACGCCATGGCTGAGGGCACGAAGGCTGTGTCGAACTCGTGCCGTTAAGTCGGTCGGTCTGGAGCCGTGGAG 649
GGGAGAGGGTTGCCGCTGCATCTGTTGTCCTTGGCGACTGCACCGCGGCCCACCCCTCACCCTCCTCTTCTTTCCCTCCCTCTTCTCTCG 739
CGGTGTGCGTCATGCAACTGCTTGACCACCCCTCTCATAGATGCAACTTTGTTTTCTTGCCCTTTTTTCTGTTTGGTTTGTAGCCGCATT 829
GACCGCTGGGCGACGTGAGCAGTG[CGCGCTGGCAGCAGTGCG]AATCTTGTCAACCACTTCTTATCGCTGTTGGTGCTACCCGCGGTCT 917
AAAAA>
GCCACGCCGCTGATGGTTCCTATCTTACCTTGTGTGCCTGCCTGCACACTCTGCGGAGAAGGGTGCCACGACGCCGGTGCCGCCTAACCT 1007
CGCCTTCTTTCGCATGTCCGCTACGCGCAGCCTTTCCTTGCCCTCCGACTACCCCGGCTCCATCCTCAACGTACCTTTCTCTCCCCACCG 1097
ACTTTCCTTCTCTTTTCTGCGCTCGCGTGCTGTTACACCTTCCCATGCTTCTCCACGCCTTCCGGACACTCCGCTCTTCTTCGCCTACCC 1187
MetAlaSerGlnArgLeuAlaProArgLysSerSerAlaAlaHisLysSerHisArgLysProLysArg
CCAATCCCCCCACGCTGCC--ATGGCCTCTCAACGTCTTGCTCCCCGCAAGTCCTCTGCGGCGCACAAGTCGCACCGCAAGCCGAAGCGC 1275
<c18
SerTrpAsnValTyrValSerArgSerLeuLysAlaIleAsnSerHisMetSerMetSerGlyArgThrMetLysIleValAsnSerTyr
TCGTGGAACGTGTACGTGAGCCGCTCTCTGAAGGCGATCAACAGCCACATGTCGATGTCTGGCCGCACGATGAAGATTGTGAACTCGTAC 1365
ValAsnAspValMetGluArgIleAlaMetGluAlaAlaSerIleValArgAlaHisLysLysLeuThrLeuGlyAlaArgGluValGln
GTGAACGACGTGATGGAGCGCATTGCGATGGAAGCCGCGTCGATTGTTCGCGCGCACAAGAAGCTCACGCTGGGCGCGCGCGAGGTGCAG 1455
ThrAlaValArgLeuValLeuProProGluLeuAlaLysHisAlaMetAlaGluGlyThrLysAlaValSerAsnSerCysArgOch
ACGGCGGTGCGCCTTGTGCTGCCGCCCGAGCTCGCGAAGCACGCCATGGCTGAGGGCACGAAGGCTGTGTCGAACTCGTGCCGTTAAGTC 1545
GGTCGGTCTGGAGCCGTGGAGGGGAGAGGGTTGCCGCTGCATCTGTTGTCCTTGGCGACTGCACCGCGGCCCACCCCTCACCCTCCTCTT 1635
CTTTCCCTCCCTCTTCTCTCGCGGTGTGCGTCATGCAACTGCTTGACCACCCCTCTCATAGATGCAACTTTGTTTTCTTGCCCTTTTTTC 1725
TGTTTGGTTTGTAGCCGCATTGACCGCTGGGCGACGTGAGCAGTG[CGCGCTGGCAGCAGTGCG]AATCTTGTCAACCACTTCTTATCGC 1813
AAAAA>
TGTTGGTGCTACCCGCGGTCTGCCACGCCGCTGATGGTTCCTATCTTACCTTGTGTGCCTGCCTGCACACTCTGCGGAGAAGGGTGCCAC 1903
GACGCCGGTGCCGCCTAACCTCGCCTTCTTTCGCATGTCCGCTACGCGCAGCCTTTCCTTGCCCTCCGACTACCCCGGCTCCATCCTCAA 1993
CGTACCTTTCTCTCCCCACCGACTTTCCTTCTCTTTTCTGCGCTCGCGTGCTGTTACACCTTCCCAGCTTCTCCACGCCTTCCAGACACT 2083
MetAlaSerGlnArgLeuAlaProArgLysSerSerAlaAlaHisLysSer
CCGCTCTTCTTCGCCTACCCCCAATCCCCCCACGCTGCCATGGCCTCTCAACGTCTTGCTCCCCGCAAGTCCTCTGCGGCGCACAAGTCG 2171
<c43
HisArgLysProLysArgSerTrpAsnValTyrValSerArgSerLeuLysAlaIleAsnSerHisMetSerMetSerGlyArgThrMet
CACCGCAAGCCGAAGCGCTCGTGGAACGTGTACGTGAGCCGCTCTCTGAAGGCGATCAACAGCCACATGTCGATGTCTGGCCGCACGATG 2261
LysIleValAsnSerTyrValAsnAspValMetGluArgIleAlaMetGluAlaAlaSerIleValArgAlaHisLysLysLeuThrLeu
AAGATTGTGAACTCGTACGTGAACGACGTGATGGAGCGCATTGCGATGGAAGCCGCGTCGATTGTTCGCGCGCACAAGAAGCTCACGCTG 2351
GlyAlaArgGluValGlnThrAlaValArgLeuValLeuProProGluLeuAlaLysHisAlaMetAlaGluGlyThrLysAlaValSer
GGCGCGCGCGAGGTGCAGACGGCGGTGCGCCTTGTGCTGCCGCCCGAGCTCGCGAAGCACGCCATGGCTGAGGGCACGAAGGCTGTGTCG 2441
AsnSerCysArgOch
AACTCGTGCCGTTAAGTCGGTCGGTCTGGAGCCGTGGAGGGGAGAGGGTTGCCGCTGCATCTGTTGTCCTTGGCGACTGCACCGCGGCCC 2531


ACCCCTCACCCTCCTCTTCTTTCCCTCCCTCTTCTCTCGCGGTGTGCGTCATGCAACTGCTTGACCACCCCCTCTCATAGACGCAGTTTTT 2621
TTTCTTTTCCCAATTTTTTTCTCCTCCCCTCTTCCCGAG[CGTGAAGTGTGCTTGACG]GCCTTCTGGACAGGAAGCAACGAAGAAATTC 2709
AAAAA>
CTCTTACCCCGCGTTACTCACGGTGGATTTGTGGCTAGGGCGATGCGTGCAACCGCTGTACGACGGTGGTGGACGAGCTTTTGCGTGACG 2799
CTGAAGCCAACCATTTGCTCTCCTTATGGGATGCTATTTCGATCCTCATGTTTATGTGGAGCACGAGCACCGGCTAAACCTCGTAGTGAT 2889
ACCGCCTAGCGAGGGACGCGTGGGGAATCGCGTGCACAACCTCATCCGGTGTCCTCCCGATCTCCTCCCAATAGTCCTTTCCCTCTCCTC 2979
ACTTCTGTACGCCTTATTCGCCCCTGCGACGATGTGCACACACACACACACACACACACACACACACACACACACACACTCGCACACGCG 3069
ATCACTGGGGCCCTTGTGCAGCTTGCTGAACACACGCGATTCTCTTTTGTACTCTTCTAACCCTCTTTTTCTTCGTGAGAAGCATTAGCC 3159
CTGACGCGAAACGAGAGAGCCACAACGCAAAGGTCAAGATCGTACAGGTCTATTTTTTTTTGAAAATGAGCGCAGGGCCGAAGCCTGAGG 3249
CCCGCCAGGCCCCGAGGGAGCGTAACACGCCGAAGCCGCCCACCGGCAATTCCAGTAGTGCTCATGCAGAGCGCCCGCAGCGCCGCCGTG 3339
ACAATGGGGAGCGCTGTGAGGGTGCGGAGCGGGGTGAGCGCCGTGATCGACCGGAGCGCCACGAGCGTCTGAGCCGTGCGCCGGCGGGTT 3429
CGGCTGCTGCCGCCGAAAAGAGCAAGCCATCTATGGAGGAAAAGGCGCATGGCCCACTTGTGCACGTGAAGGGCGACATTGCGTCGGGG- 3518
GTCGAC-Sal-1 3524

Figure 11 Densitometry analysis of *Pst* I digested genomic DNA. Five µg of *Pst* I digested *Leishmania enriettii* DNA was separated on a 1% agarose gel, blotted and hybridized with radiolabelled c18 insert and exposed to pre-fogged film. Autoradiograms were scanned with the densitometer, and optical densities beneath peaks were automatically integrated. Peaks A, B, and C correspond to bands A, B, and C of the Southern (inset).

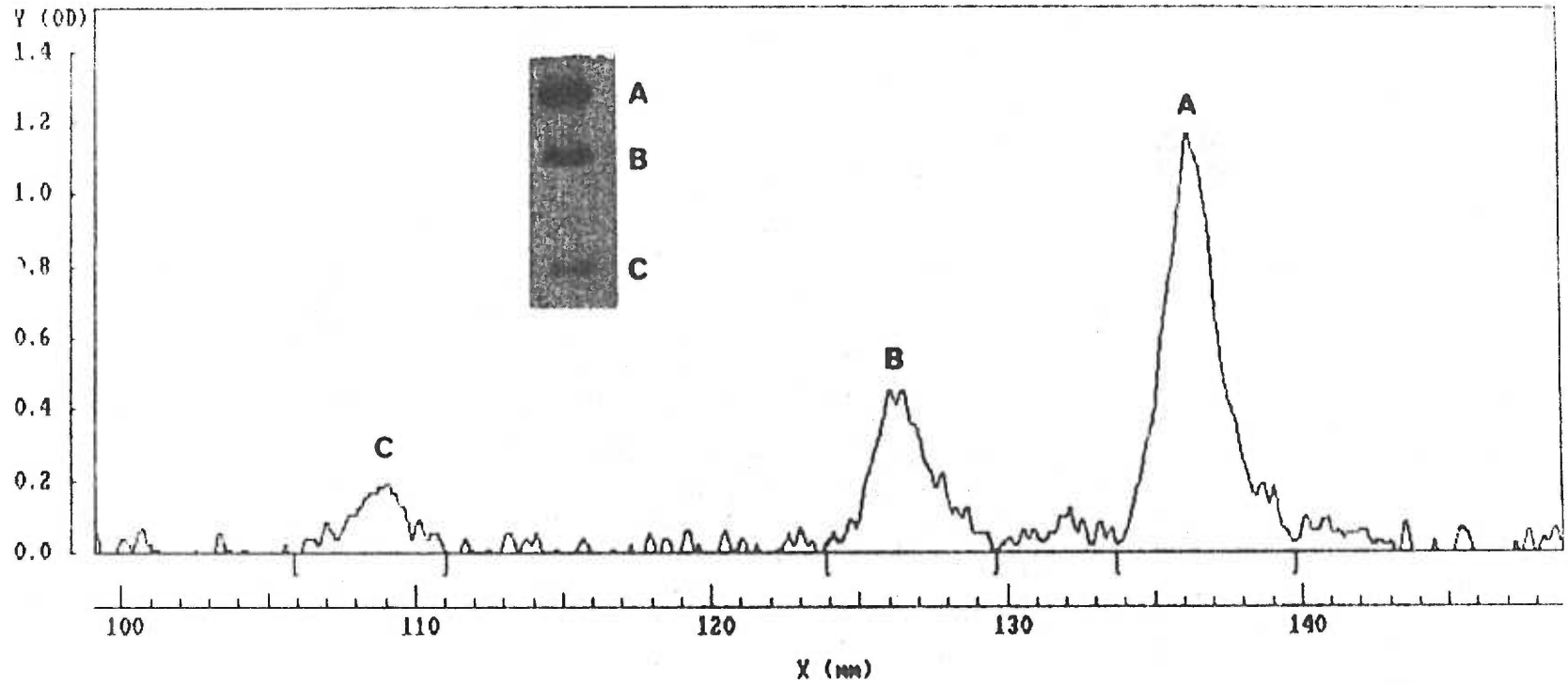

Y (OD)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
A
B
C
A
B
C
100
110
120
130
140
X (mm)

Figure 12 Chromosomal analysis of *Leishmania enriettii* histone H2B genes. Chromosomes from *L. enriettii* were separated on a 1.5% agarose gel using OFAGE electrophoresis. Gels were stained with ethidium bromide and photographed (lanes 1 and 2 on right), and transferred to filters for hybridization (left). Filters were cut into lanes and hybridized with insert from c18 (lanes 1) and insert from Pro-1 transporter clone (lanes 2). The star, circle and arrowhead indicate the postions of hybridizing chromosomes (left) in the ethidium bromide stained gel (right).

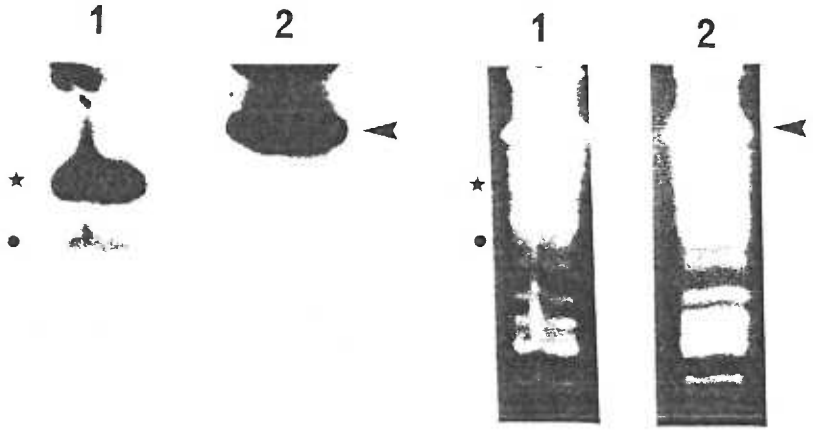
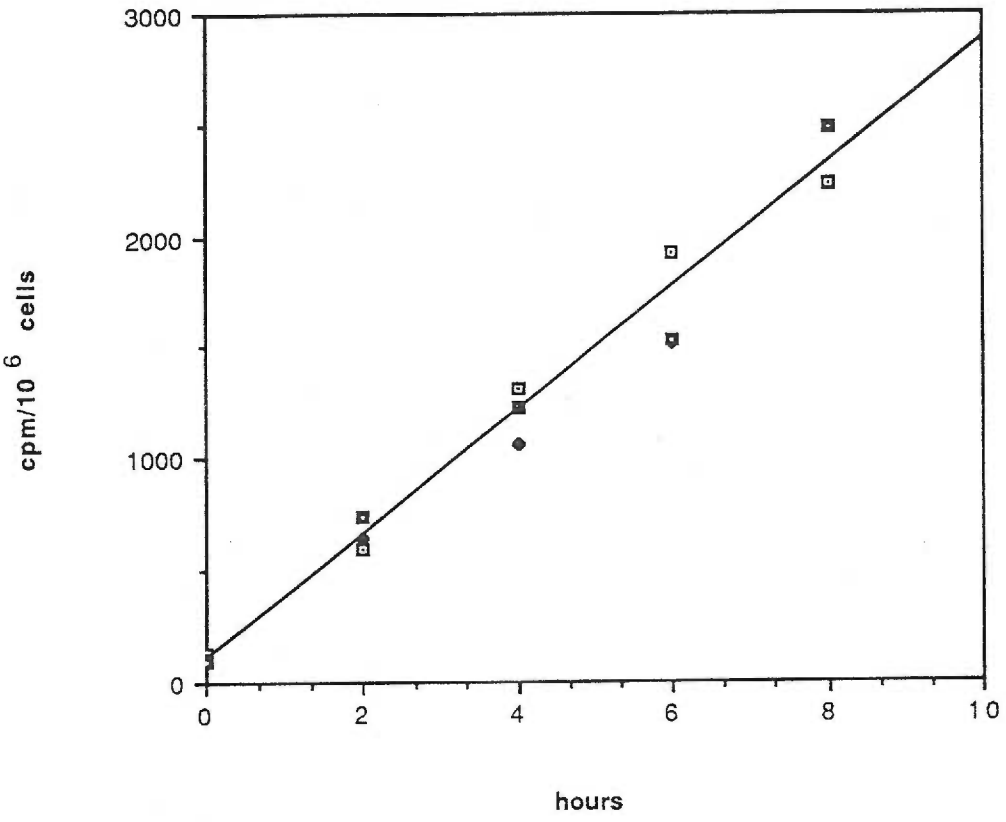

1
2
1
2

Figure 13 (A) Incorporation of ^{3}H-hypoxanthine during DNA synthesis. Promastigotes were grown in the presence of ^{3}H-hypoxanthine and harvested in triplicate at two hour intervals for 8 hours. Cells were counted using a hemocytometer, washed in PBS and incubated overnight in 0.3 M KOH. DNA was TCA-PPi precipitated and counted. (B) Same conditions, except cells were treated with DNA synthesis inhibitor 5 mM hydroxyurea or 10 μM aphidicolin at time 0.

A



B

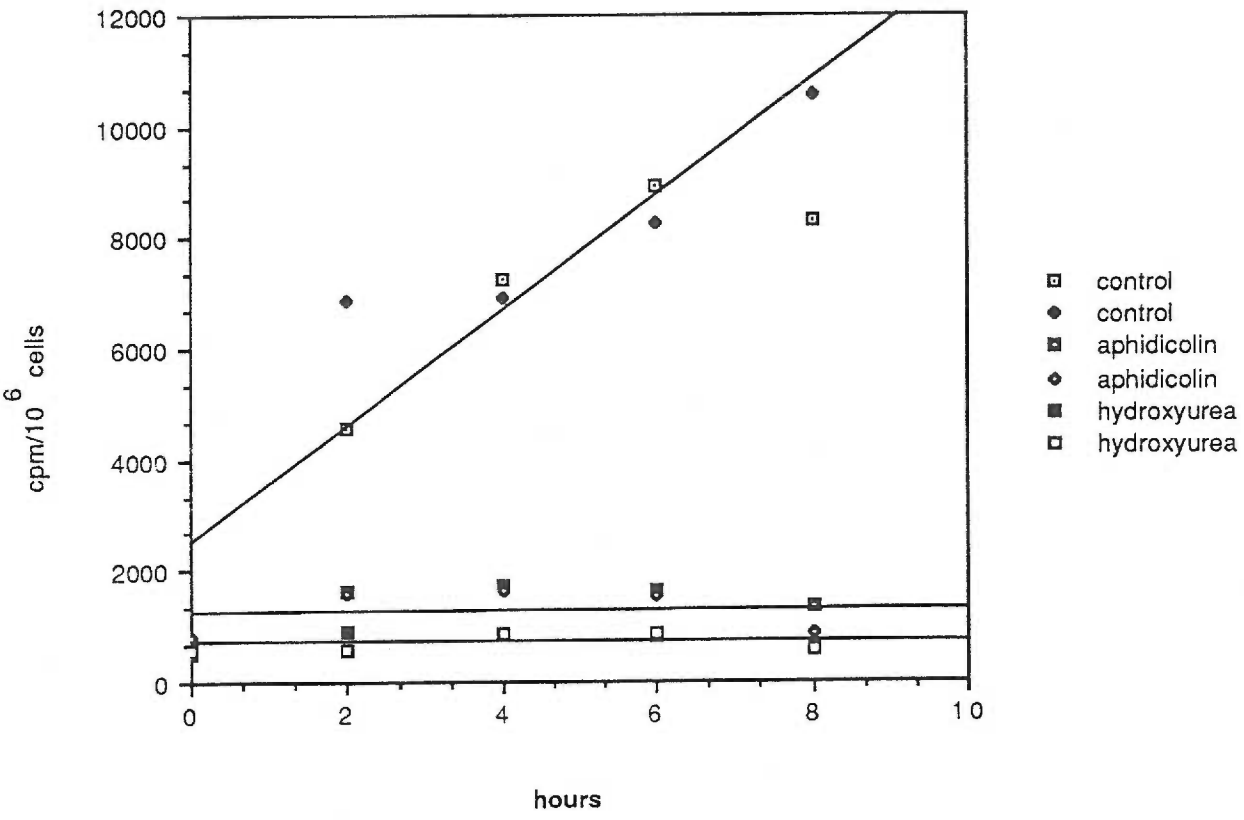
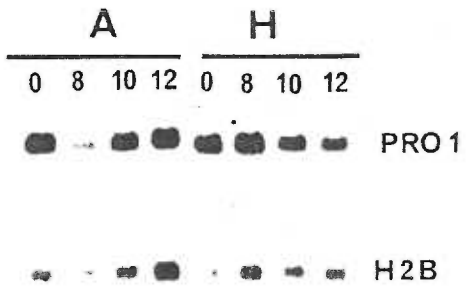
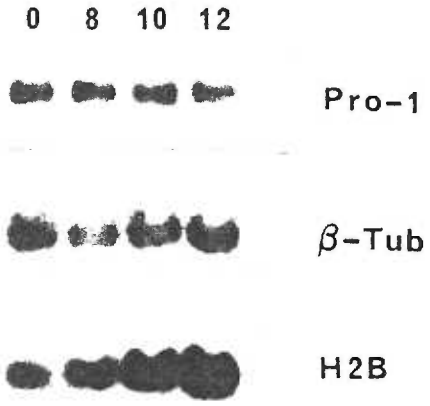

12000
10000
8000
6000
4000
2000
0
cpm/10 6 cells
0
2
4
6
8
10
hours
control
control
aphidicolin
aphidicolin
hydroxyurea
hydroxyurea

Figure 14 A. Aphidicolin treated and hydroxyurea treated cells. DNA synthesis inhibitors hydroxyurea or aphidicolin were added to exponentially growing cells. Total RNA was harvested at time 0, 8 hr, 10 hr, and 12 hr. 5 μg of RNA from each sample was separated on a 1% agarose-formaldehyde gel. The gel was transferred to nitrocellulose filters and hybridized to a mixture of the radiolabelled inserts c18 and Pro-1. A = aphidicolin treated RNA, H = hydroxyurea treated RNA. B. 5 μg of total RNA from cells treated with aphidicolin for 0 hr., 8 hr., 10 hr., and 12 hr., was separated and transferred as above. The filter was hybridized with a mixture of radiolabelled inserts from c18 histone H2B cDNA clone, β-tubulin genomic clone (33), and Pro-1 transporter genomic clone (9). (C) 5 μg of total RNA from cells incubated with hydroxyurea for 0 hr, 8 hr, 10 hr, and 12 hr separated under same conditions as A, probed with c18.


A
H
0 8 10 12 0 8 10 12
PRO 1
H2B

B



C

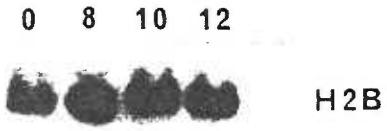


Figure 15 S1 mapping of the genomic clone P6.4. Linearized P6.4 was denatured at 75°C and annealed to total promastigote RNA at 62°C. S1 nuclease (300 units) was added to the reaction and incubated at 45°C. Phenol/chloroform extracted and ethanol precipitated samples were separated on 1% alkaline agarose gel and transferred to a nitrocellulose filter. Filters were hybridized with radiolabelled insert from histone H2B cDNA c61-1. Lane A, DNA alone; lane B, DNA and S1 nuclease; lane C, DNA, RNA, and S1 nuclease. ϕX174 *Hae* III digested DNA is shown on the right with the molecular weight sizes (MW). The arrows indicate hybridizing bands of greater molecular weight than the full length mRNA (750 bp).

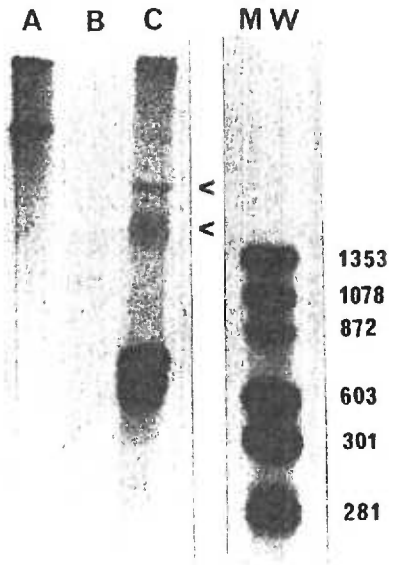

A
B
C
M W
1353
1078
872
603
301
281

Figure 16 Schematic representation of histone H2B mRNA structures. Cell cycle regulated histone mRNAs from higher eukaryotes (top) have a conserved structure. They do not contain intervening sequences, or poly-A tails, but instead have a highly conserved hairpin structure at the 3' terminus. Yeast histones and variant histones (middle) differ in mRNA structure from the cell cycle regulated histones in that they are both polyadenylated. In addition variant histones contain intervening sequences. The *Leishmania* histone H2B mRNA identified here (bottom) is different from both the higher eukaryote cell cycle regulated histones and the yeast and variant histones. It is unique structurally in containing both a poly-A tail and a novel hairpin loop at the 3' terminus.

Cell Cycle
Regulated
Histones

Variant Histones
and
Yeast Histones

Leishmania
Histone H2b