The EMBO Journal vol.9 no.4 pp.1027–1033, 1990

The amplified H circle of methotrexate-resistant *Leishmania tarentolae* contains a novel P-glycoprotein gene

Marc Ouellette, Francesca Fase-Fowler and Piet Borst

Division of Molecular Biology, Netherlands Cancer Institute, Antoni van Leeuwenhoekhuis, Plesmanlaan 121, 1066CX Amsterdam, The Netherlands

Communicated by P.Borst

Acquired resistance to methotrexate in *Leishmania* species is often associated with the amplification of H circles, 68 kb duplex DNA circles containing a 30 kb inverted repeat. We report here that the H circle of *Leishmania tarentolae* contains an open reading frame, *ltpgpA*, that has the attributes of P-glycoproteins (large plasma membrane proteins known to extrude lipophilic drugs from mammalian cells). Although amplification of H circles is associated with proportionally increased levels of a 5.5 kb transcript of the *ltpgpA* gene, such methotrexate resistant mutants are not cross-resistant to any of the drugs extruded by mammalian multi-drug resistant cells. In *Leishmania, ltpgpA* is part of a gene family containing at least two other members. Sequences homologous to one of the nucleotide binding sites of ltpgpA are conserved in other kinetoplastida.

Key words: gene amplification/*Leishmania*/methotrexate resistance/P-glycoprotein/plasmid

Introduction

Acquired resistance to methotrexate (MTX) is readily induced in the kinetoplastid protozoan flagellate *Leishmania* under laboratory conditions. Resistance may involve decreased uptake of MTX (Ellenberger and Beverley, 1987; Kaur *et al.*, 1988), amplification of the gene for the bifunctional dihydrofolate reductase–thymidylate synthase (Beverley *et al.*, 1984), or by other mechanisms that remain to be defined.

In some resistant *Leishmania* mutants, resistance correlates with amplification of H circles (Hightower *et al.*, 1988; Petrillo-Peioxoto and Beverley, 1988; White *et al.*, 1988). These extrachromosomal circles contain a 30 kb inverted repeat, separated by 4 and 5 kb unique sequences (Figure 1). There is also a chromosomal copy of the H circle sequence, but without the inverted repeat. We have shown that some wild type strains of *Leishmania tarentolae* contain H circles whereas others do not (White *et al.*, 1988). Induction of MTX resistance is often associated with amplification of pre-existing H circles (White *et al.*, 1988), or *de novo* generation of H circles from the chromosomal copy (unpublished results).

Although there is a clear link between the 68 kb H circle amplification and MTX resistance (Hightower *et al.*, 1988; Petrillo-Peioxoto and Beverley, 1988; White *et al.*, 1988) the nature of this link is unclear. Amplification of the H circle has also been observed in *Leishmania major* cell lines selected with terbinafine and primaquine, two drugs that are structurally and mechanistically unrelated to MTX (Ellenberger and Beverley, 1989). A 69 kb circle was amplified as well in a *Leishmania mexicana amazonensis* cell line made resistant to arsenite. This variant was cross resistant to MTX and its amplified circle hybridized with an H circle probe (Katakura and Chang, 1989). Taken together these data raise the possibility that the H circle encodes a protein involved in an unusual form of multidrug resistance. In mammalian cells, resistance to a diverse set of hydrophobic drugs can be caused by increased levels of P-glycoproteins, large plasma membrane proteins, that function as ATP dependent drug extrusion pumps (see Gottesman and Pastan, 1988; Endicott and Ling, 1989; van der Bliek and Borst, 1989). Such P-glycoproteins have also been implicated in chloroquine resistance in the malaria parasite *Plasmodium falciparum* (Foote *et al.*, 1989; Wilson *et al.*, 1989). We have therefore tested whether the H-circles of *L.tarentolae* encode a P-glycoprotein.

Results

H-circles contain a P-glycoprotein gene

The P-glycoprotein genes analysed thus far share highly conserved sequences covering the nucleotide binding domains in the protein (Endicott and Ling, 1989; van der Bliek and Borst, 1989). We have used a cDNA probe containing this area from the human *mdr1* gene (Chen *et al.*, 1986) to search for homologous sequences in purified H circles. A single segment of weak hybridization was found

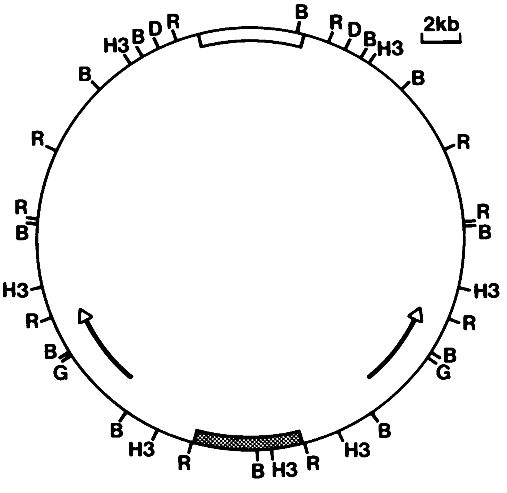


Fig. 1. Location of the P-glycoprotein gene A on the H circle of *L.tarentolae*. The structure and map of the H circle is taken from White *et al.* (1988). The H circle is 68 kb with 30 kb inverted repeats (thin line) with two unique regions (open and dotted rectangles). Arrows inside the circle indicate the location of the P-glycoprotein gene. B, *Bam*HI; D, *Dra*I; R, *Eco*RI; G, *Bgl*II; H3, *Hin*dIII.

© Oxford University Press

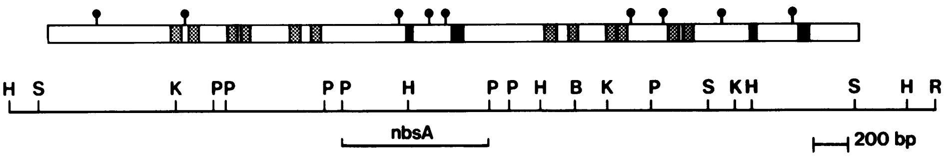


Fig. 2. Restriction map of the *L.tarentolae* P-glycoprotein gene A and a schematic representation of the corresponding protein. The dotted rectangles correspond to putative transmembrane domains, the filled rectangles to the nbs consensus sequences, and the bars with circle to putative N-glycosylation sites. The *Pst*I fragment nbsA was used in the experiments shown in Figures 5 and 6. B, *Bam*HI, H, *Hin*dII, K, *Kpn*I, P, *Pst*I, R, *Eco*RI and S, *Sac*I.

```
1      ACCTCTGAACTGTACTAGCCTATGGGGAGGTGGGCGCTATAGAGATCGACTCATCGAGGGGGCATCTACAGCGCGTAGTGCATGACAACACCAGACTGTGTACGCGTGTAACCGATTGAC

                                                            M  V  D  N  G  H  V  T  I  A  M  A  D  L  G  T  V  V  E  I  A  Q  V    23
121    GATCCCAGTGAGCTCAGAATTTGTGCTCCCTCGCAGCGGCGCCACATCCAATATGGTAGACAACGGTCACGTCACCATCGCGATGGCTGATTTGGGGACAGTGGTCGAGATCGCACAGGT

         R  C  Q  Q  E  A  Q  R  K  F  A  E  Q  L  D  E  L  W  G  G  E  P  A  Y  T  P  T  V  E  D  Q  A  S  W  F  Q  Q  L  Y  Y    63
241    GCGCTGCCAGCAGGAGGCGCAACGCAAGTTCGCGGAGCAGCTAGATGAGCTGTGGGGCGGCGAACCAGCGTACACACCGACGGTAGAGGACCAGGCGAGCTGGTTTCAGCAGCTCTACTA

         G  W  I  G  D  Y  I  Y  K  A  A  A  G  N  I  T  E  A  D  L  P  P  P  T  R  S  T  R  T  Y  H  I  G  R  K  L  S  R  Q  A    103
361    CGGGTGGATTGGGGACTACATTTACAAGGCTGCTGCGGGAAACATCACCGAGGCGGACCTGCCACCGCCGACGCGGAGCACACGGACTTACCACATAGGGCGCAAGTTGTCACGGCAGGC

         H  A  D  I  D  A  S  R  R  W  Q  G  Y  I  G  C  E  V  V  Y  K  S  E  A  E  A  K  G  V  L  R  W  V  G  H  L  Q  Q  S  D    143
481    GCATGCCGACATCGACGCAAGTCGGCGGTGGCAGGGGTACATCGGGTGCGAGGTCGTGTACAAGTCAGAGGCAGAAGCGAAGGGCGTGCTGCGGTGGGTTGGGCATTTGCAGCAGTCAGA

         Y  P  R  S  L  V  A  G  V  E  W  R  M  P  P  R  H  R  R  L  A  V  L  G  S  A  A  A  L  H  N  G  V  V  H  G  E  R  L  F    183
601    CTACCCGCGATCGCTGGTCGCTGGGGTAGAGTGGCGCATGCCGCCTCGGCACCGGCGGCTGGCCGTGCTGGGCAGTGCGGCAGCACTTCACAACGGCGTTGTGCATGGTGAGCGTCTGTT

         W  P  H  E  D  N  Y  L  C  S  C  E  P  V  E  Q  L  Y  V  K  S  K  Y  N  L  I  P  P  R  P  P  P  S  P  D  L  L  R  T  L    223
721    TTGGCCACATGAAGACAACTATCTCTGCTCGTGCGAGCCGGTGGAGCAGTTGTATGTAAAGTCGAAGTACAACCTGATTCCTCCCCGGCCGCCGCCGTCGCCGGATCTTCTGCGGACGTT

         F  K  V  H  W  Y  H  V  W  A  Q  I  L  P  K  L  L  S  D  V  T  A  L  M  L  P  V  L  L  E  Y  F  V  K  Y  L  N  A  D  N    263
841    GTTCAAGGTGCACTGGTACCACGTGTGGGCGCAGATCCTGCCAAAGCTGTTGTCGGATGTTACCGCGTTGATGCTGCCGGTGCTGTTGGAGTACTTTGTGAAATATCTGAACGCCGATAA

         A  T  W  G  W  G  L  G  L  A  L  T  I  F  L  T  N  V  I  Q  S  C  S  A  H  K  Y  D  H  I  S  I  R  T  A  A  L  F  E  T    303
961    CGCGACGTGGGGGTGGGGTCTGGGACTTGCGCTCACAATCTTCCTTACGAACGTAATTCAGAGCTGCTCAGCGCACAAGTATGACCACATCAGCATCCGCACTGCAGCGCTGTTTGAGAC

         S  S  M  A  L  L  F  E  K  C  F  T  V  S  R  R  S  L  Q  R  P  D  M  S  V  G  R  I  M  N  M  V  G  N  D  V  D  N  I  G    343
1081   GTCATCGATGGCGCTTCTGTTCGAGAAGTGCTTCACGGTCTCGCGGCGATCACTGCAGCGCCCGGACATGTCCGTGGGTCGCATCATGAACATGGTTGGGAATGACGTTGATAACATCGG

         S  L  N  W  Y  V  M  Y  F  W  S  A  P  L  Q  L  V  L  C  L  L  L  L  I  R  L  V  G  W  L  R  V  P  G  M  A  V  L  F  V    383
1201   TAGCCTTAACTGGTACGTGATGTACTTTTGGAGCGCCCCACTGCAACTGGTGCTATGCCTTCTGCTGCTCATAAGGCTGGTCGGGTGGCTTCGCGTGCCCGGCATGGCCGTCTTGTTCGT

         T  L  P  L  Q  A  V  I  S  K  H  V  Q  D  V  S  E  R  M  A  S  V  V  D  L  R  I  K  R  T  N  E  L  L  S  G  V  R  I  V    423
1321   GACGCTCCCATTGCAAGCCGTCATTTCCAAGCACGTTCAGGACGTGTCTGAGCGAATGGCGAGCGTGGTGGACCTCCGCATCAAGCGCACGAACGAGCTTCTCTCCGGTGTTCGAATCGT

         K  F  M  G  W  E  P  V  F  L  A  R  I  Q  D  A  R  S  R  E  L  R  C  L  R  D  V  H  V  A  N  V  F  F  M  F  V  N  D  A    483
1441   GAAATTCATGGGCTGGGAGCCTGTCTTTCTCGCCCGTATCCAAGACGCGCGCAGTCGTGAGCTTCGGTGTTTGCGAGATGTGCATGTGGCTAATGTATTCTTTATGTTCGTGAATGATGC

         T  P  T  L  V  I  A  V  V  F  I  L  Y  H  V  S  G  K  V  L  K  P  E  V  V  F  P  T  I  A  L  L  N  T  M  R  V  S  F  F    523
1561   AACACCGACGTTGGTAATTGCGGTCGTCTTCATTTTGTACCATGTAAGTGGGAAGGTGCTGAAGCCAGAGGTTGTGTTCCCAACGATTGCCCTCCTCAACACTATGCGCGTGTCATTCTT

         M  I  P  I  I  I  S  S  I  L  Q  C  F  V  S  A  K  R  V  T  A  F  I  E  C  P  D  T  H  S  Q  V  Q  D  I  A  S  I  D  V    563
1681   CATGATACCAATTATCATATCTTCCATTCTGCAGTGCTTTGTGTCGGCAAAGCGTGTGACGGCTTTTATAGAGTGCCCAGACACGCACTCACAGGTGCAAGACATTGCTAGCATTGACGT

         P  D  A  A  A  I  F  K  G  A  S  I  H  T  Y  L  P  V  K  L  P  R  C  K  S  R  L  T  A  M  Q  R  S  T  L  W  F  R  R  R    603
1801   GCCGGATGCTGCAGCCATCTTCAAGGGTGCGTCCATCCATACGTATTTGCCGGTGAAACTCCCTCGATGCAAGTCGCGCTTGACAGCCATGCAGCGCAGTACGCTGTGGTTCCGCCGGCG

         G  V  P  E  T  E  W  Y  E  V  D  S  P  D  A  S  A  S  S  L  A  V  H  S  T  T  V  H  M  G  S  T  Q  T  V  I  T  D  S  D    643
1921   CGGTGTGCCGGAGACAGAGTGGTACGAGGTGGACAGCCCCGATGCGAGTGCGTCTTCACTGGCAGTACACAGCACAACGGTCCACATGGGCAGCACACAGACCGTCATTACAGACAGCGA

         G  A  A  G  E  D  E  K  G  E  V  E  E  G  D  R  E  Y  Y  Q  L  V  S  K  E  L  L  R  N  V  S  L  T  I  P  K  G  K  L  T    683
2041   TGGGGCTGCTGGTGAAGACGAGAAGGGAGAAGTGGAAGAGGGAGACAGAGAGTACTACCAACTTGTGTCAAAGGAACTGCTACGGAACGTGAGCCTGACCATCCCGAAAGGGAAGCTGAC

         M  V  I  G  S  T  G  S  G  K  S  T  L  L  G  A  L  M  G  E  Y  S  V  E  S  G  E  L  W  A  E  R  S  I  A  Y  V  P  Q  Q    723
2161   AATGGTGATTGGCTCGACAGGGAGCGGGAAGTCAACACTTCTGGGTGCGCTGATGGGCGAGTACAGCGTAGAGAGCGGGGAGCTGTGGGCAGAGCGGAGCATTGCGTACGTGCCGCAGCA

         A  W  I  M  N  A  T  L  R  G  N  I  L  F  F  D  E  E  R  A  E  D  L  Q  D  V  I  R  C  C  Q  L  E  A  D  L  A  Q  F  C    763
2281   GGCGTGGATCATGAACGCGACGCTGCGTGGCAATATTCTGTTCTTCGACGAGGAGCGTGCCGAGGACTTGCAGGATGTGATACGGTGCTGCCAGTTGGAGGCGGACCTTGCGCAGTTTTG

         G  G  L  D  T  E  I  G  E  M  G  V  N  L  S  G  G  Q  K  A  R  V  S  L  A  R  A  V  Y  A  N  R  D  V  Y  L  L  D  D  P    803
2401   CGGTGGGCTGGACACGGAGATTGGGGAGATGGGCGTAAACCTGAGCGGCGGGCAAAAGGCACGCGTGAGCCTCGCACGCGCCGTGTACGCGAACCGCGACGTGTACCTGCTGGACGACCC

         L  S  A  L  D  A  H  V  G  Q  R  I  V  Q  D  V  I  L  G  R  L  R  G  K  T  R  V  L  A  T  H  Q  I  H  L  L  P  L  A  D    843
2521   CCTGTCCGCGCTGGACGCGCACGTTGGCCAGCGCATCGTGCAAGACGTCATCCTTGGACGGCTGCGCGGCAAGACGCGCGTGCTTGCAACGCACCAGATACATCTACTGCCGCTTGCAGA

         Y  I  V  V  L  Q  H  G  S  I  V  F  A  G  D  F  A  A  F  S  A  T  A  L  E  E  T  L  R  G  E  L  K  G  S  K  D  V  E  S    883
2641   TTACATTGTGGTACTGCAGCATGGCAGCATCGTGTTTGCTGGCGACTTTGCCGCTTTTTCGGCAACTGCCCTGGAGGAGACGCTGCGTGGTGAGCTAAAGGGGAGCAAGGATGTGGAGTC

         C  S  S  D  V  D  T  E  S  A  T  A  E  T  A  P  Y  V  A  K  A  K  G  L  N  A  E  Q  E  T  S  L  A  G  G  E  D  P  L  R    923
2761   CTGCAGCAGCGATGTGGACACAGAGTCAGCAACAGCTGAAACGGCACCGTATGTTGCCAAAGCGAAAGGTCTGAATGCTGAGCAAGAGACGAGCCTTGCAGGGGGCGAAGATCCCTTGAG

         S  D  V  E  A  G  R  L  M  T  T  E  E  K  A  T  G  K  V  P  W  S  T  Y  V  A  Y  L  K  S  C  G  G  L  E  A  W  G  C  L    963
2881   GTCTGATGTCGAGGCCGGGAGGCTGATGACCACGGAGGAGAAGGCAACCGGTAAGGTGCCGTGGTCAACCTACGTCGCTTATCTGAAGTCGTGTGGCGGCCTGGAAGCGTGGGGGTGCCT

         L  A  T  F  A  L  T  E  C  V  T  A  A  S  S  V  W  L  S  I  W  S  T  G  S  L  M  W  S  A  D  T  Y  L  Y  V  Y  L  F  I    1003
3001   GCTAGCGACCTTTGCGCTAACAGAGTGTGTGACCGCAGCTAGCAGCGTGTGGCTGTCGATATGGTCCACCGGCTCGCTGATGTGGAGCGCGGACACGTACCTCTACGTCTATCTTTTCAT

         V  F  L  E  I  F  G  S  P  L  R  F  F  L  C  Y  Y  L  I  R  I  G  S  R  N  M  H  R  D  L  L  E  S  I  G  V  A  R  M  S    1043
3121   TGTTTTCCTTGAGATCTTCGGATCCCCACTGCGATTTTTCTTGTGCTACTACCTTATTCGCATTGGCAGCCGAAACATGCATCGCGACCTTCTCGAGTCTATTGGAGTTGCGCGGATGTC

         F  F  D  T  T  P  V  G  R  V  L  N  R  F  T  K  D  M  S  I  L  D  N  T  L  N  D  G  Y  L  Y  L  L  E  Y  F  F  S  M  C    1083
3241   TTTCTTTGATACAACCCCCGTAGGACGCGTACTCAACCGCTTTACCAAGGACATGAGCATCCTCGACAACACGTTGAATGACGGGTACCTTTACCTGCTGGAGTACTTTTTTTCCATGTG

         S  T  V  I  I  M  V  V  V  Q  P  F  V  L  V  A  I  V  P  C  V  Y  S  Y  Y  K  L  M  Q  V  Y  N  A  S  N  R  E  T  R  R    1123
3361   CTCCACCGTCATTATCATGGTGGTTGTGCAACCCTTTGTTCTGGTGGCCATTGTGCCGTGCGTGTACAGCTACTACAAGCTGATGCAGGTCTACAACGCATCGAACCGCGAGACACGCCG

         I  K  S  I  A  H  S  P  V  F  T  L  L  E  E  S  L  Q  G  Q  R  T  I  A  T  Y  G  K  L  H  L  V  L  Q  E  A  L  G  R  L    1163
3481   CATCAAAAGTATCGCACACTCGCCAGTATTCACGCTGCTGGAGGAGTCGCTCCAAGGGCAGCGCACCATCGCGACGTACGGCAAGTTACATCTCGTGCTGCAGGAGGCATTGGGGCGGTT
```

```
           D  V  V  Y  S  A  L  Y  M  Q  N  V  S  N  R  W  L  G  V  R  L  E  F  L  S  C  V  V  T  F  M  V  A  F  I  G  V  I  G  K    1703
3601   GGATGTGGTGTATAGCGCTCTGTATATGCAGAACGTCTCCAACCGTTGGCTCGGTGTCCGTCTCGAGTTTTTGAGCTGCGTTGTCACCTTTATGGTGGCCTTCATTGGCGTGATAGGGAA

           M  E  G  A  S  S  Q  N  I  G  L  I  S  L  S  L  T  M  S  M  T  L  T  E  T  L  N  W  L  V  R  Q  V  A  M  V  E  A  N  M    1243
3721   GATGGAGGGTGCATCGAGCCAGAACATCGGCCTCATATCGCTTTCACTAACCATGTCGATGACGCTGACAGAGACACTCAACTGGCTGGTGCGCCAAGTTGCCATGGTGGAAGCTAACAT

           N  S  V  E  R  V  L  H  Y  T  Q  E  V  E  H  E  H  V  P  E  M  G  E  L  V  A  Q  L  V  R  S  E  S  G  R  G  A  N  V  T    1283
3841   GAACAGCGTGGAACGTGTGCTACACTACACTCAGGAAGTGGAACACGAGCACGTTCCGGAGATGGGGGAGCTCGTGGCACAGCTGGTAAGGAGCGAGTCAGGGAGAGGGGCTAACGTGAC

           E  T  V  V  I  E  S  A  G  A  A  S  S  A  L  H  P  V  Q  A  G  S  L  V  L  E  G  V  Q  M  R  Y  R  E  G  L  P  L  V  L    1323
3961   CGAGACGGTTGTGATCGAGTCCGCGGGCGCGGCGAGCAGCGCACTGCACCCGGTGCAGGCCGGGTCGCTTGTGCTCGAGGGTGTGCAGATGCGGTACCGGGAGGGGCTGCCGCTTGTGCT

           R  G  V  S  F  Q  I  A  P  R  E  K  V  G  I  V  G  R  T  G  S  G  K  S  T  L  L  L  T  F  M  R  M  V  E  V  C  G  G  V    1363
4081   GCGCGGCGTGAGCTTCCAGATCGCGCCGCGCGAGAAGGTCGGCATTGTCGGGCGGACAGGGAGTGGGAAGTCGACGCTGCTGCTGACGTTCATGCGGATGGTTGAGGTGTGCGGCGGGGT

           I  H  V  N  G  R  E  M  S  A  Y  G  L  R  E  L  R  R  H  F  S  M  I  P  Q  D  P  V  L  F  D  G  T  V  R  Q  N  V  D  P    1403
4201   GATCCACGTGAACGGGCGCGAGATGAGTGCGTACGGGCTGCGCGAGCTGCGGCGGCACTTCTCGATGATCCCGCAGGACCCTGTGCTGTTCGACGGCACTGTGCGGCAGAACGTGGACCC

           F  L  E  A  S  S  A  E  V  W  A  A  L  E  L  V  G  L  R  E  R  V  A  S  E  S  E  G  I  D  S  R  V  L  E  G  G  S  N  Y    1443
4321   TTCCTGGAGGCGTCGTCCGCGGAGGTTTGGGCTGCGCTGGAGCTTGTGGGGCTGCGCGAGCGCGTTGCGTCGGAGAGCGAGGGTATCGACAGCCGCGTGCTGGAGGGCGGGTCGAACTA

           S  V  G  Q  R  Q  L  M  C  M  A  R  A  L  L  K  R  G  S  G  F  I  L  M  D  E  A  T  A  N  I  D  P  A  L  D  R  Q  I  Q    1483
4441   CAGCGTTGGGCAGCGGCAGCTGATGTGCATGGCGCGCGCGCTGCTGAAGAGGGGGAGCGGGTTCATCCTGATGGACGAGGCGACGGCGAACATCGACCCGGCGCTTGACCGGCAGATCCA

           A  T  V  M  S  A  F  S  A  Y  T  V  I  T  I  A  H  R  L  H  T  V  A  Q  Y  D  K  I  I  V  M  D  H  G  V  V  A  E  M  G    1503
4561   GGCGACGGTGATGAGCGCGTTCTCGGCGTACACGGTGATCACGATCGCGCACCGGCTGCACACTGTCGCGCAGTACGACAAGATCATCGTGATGGACCACGGTGTTGTTGCCGAGATGGG

           S  P  R  E  L  V  M  N  H  Q  S  M  F  H  S  M  V  E  S  L  G  S  R  G  S  K  D  F  Y  E  L  L  M  G  R  R  I  V  Q  P    1543
4681   CAGCCCGCGCGAGCTTGTGATGAACCACCAGTCCATGTTCCACAGCATGGTGGAATCGCTGGGCAGCCGTGGTTCAAAAGACTTTTATGAGCTCCTTATGGGACGTCGGATTGTGCAGCC

           A  V  L  S  D  *                                                                                                              1548
4801   GGCCGTGCTGTCTGATTAACGACTCGATTGCCACCTGAAAATCGTGAAGCAGAAGCCGCTCGCCTCGCACCACCCTGGATGCAGACGATGGGGAGAAAGCGTGGCACATCTGTGGTGCCT

4921   GGCATATTGCTGTTTTGTCTTATTTACAGGCTCATGCTTGCAGGAGGGCGGTAGCCCCGCCGCTGTCGTACTCGGGGAGGAGGTCGCATAGTGGTGGTGTGCCAGTGTGTAGAAGCACGT

5041   ATTACATCCCTAGCGGCTCTTAATTGAATCTTTATTTATGTGGAAAGCCAAGTCTTGCACAAGAGAGGTTGTCAGGGCAACACCTCGGAAGCATGTGTGAGAATGTTGAC
```

Fig. 3. Nucleotide sequence of the *L.tarentolae* P-glycoprotein gene A and deduced amino acid sequence. The nucleotide sequence starts at the first *Hin*dII site and finishes at the *Hin*dII site nearest to the *Eco*RI site of Figure 2. Numbers at the left are for nucleotides and at the right for amino acids. Above the nucleotide sequence, the deduced amino acid sequence is indicated in the one letter code. The amino acid sequences which fit the consensus for nbs of transport proteins (Higgins *et al.*, 1986) are boxed, the 12 putative transmembrane domains are underlined by a thick line and putative N-glycosylation sites are underlined by a thin line. This sequence will appear in sequence databases under the accession number X17154.

on Southern blots (marked nbsA in Figure 2). In blots of chromosome sized DNA, size-fractionated by traverse alternating field electrophoresis (see White *et al.*, 1988), the hybridizing segment was found to co-migrate with H circles. Sequence analysis showed that nbsA is part of a large open reading frame located in the inverted duplication of the H circle at the position of the arrows in Figure 1. The restriction map of this putative P-glycoprotein gene, named *ltpgpA*, is presented in Figure 2 and its nucleotide sequence in Figure 3. The longest reading frame of *ltpgpA* begins with an ATG codon at position 173 and terminates with a TAA codon at nucleotide 4819. Alternatively a second in-frame ATG at position 203 could be used. The surroundings of both proposed ATG initiation codons are in a favourable context for translation initiation (Kozak, 1987). An open reading frame initiating at nucleotide 173 would encode a protein of 1548 amino acids with a predicted mass of 172 kd. This protein is substantially larger than the P-glycoproteins found in mammalian cells, which are ~1280 amino acids (Chen *et al.*, 1986; Gerlach *et al.*, 1986; Gros *et al.*, 1986, 1988; van der Bliek *et al.*, 1988).

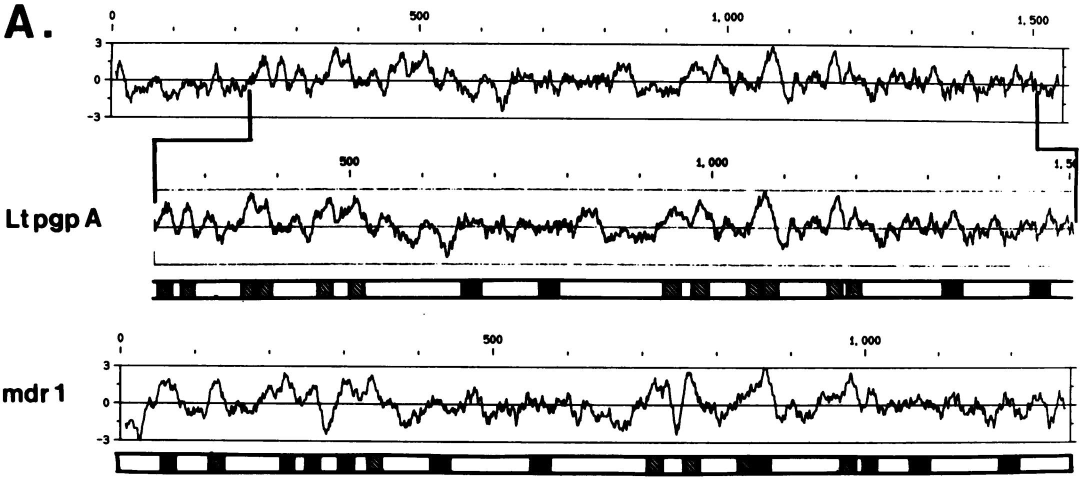
The hydropathy plots shown in Figure 4A indicate that the *ltpgpA* product consists of two similar halves, each containing six putative transmembrane segments (underlined by a thick line in Figure 3) and a nucleotide binding site (nbs), just like the human *mdr1* encoded P-glycoprotein. Figure 4B shows that the two putative nbs (boxed in Figure 3) of the ltpgpA protein are homologous to those of other P-glycoproteins. These conserved nbs motifs are found in several bacterial and eukaryotic membrane proteins, but in duplicated form only in P-glycoproteins and in *rbsA* (Table I). The homology extends to ~200 amino acids around the nbs and is not significant in the remainder of the protein or between the two halves of ltpgpA.

The additional amino acids in ltpgpA relative to other P-glycoproteins are present in an N-terminal extension. This extension neither bears homology with any protein present in data banks (Genbank release 60.0, EMBL-modified release 19.0, NBRF-protein 21.0), nor does it contain a typical signal peptide for entry into the endoplasmic reticulum (Von Heine, 1985). In fact, the N-terminal region of the extension found in ltpgpA is charged, which is usually the case for proteins that have to span a membrane several times (Friedlander and Blobel, 1985). Possible N-glycosylation sites are underlined with a thin line in Figure 3. In the predicted structure of ltpgpA only one of the N-linked glycosylation sites, located between transmembrane segments 1 and 2, would be extracellular. It is also between these first two transmembrane domains that the only extracellular N-glycosylation sites are present in mammalian P-glycoproteins and in the yeast STE6 P-glycoprotein (van der Bliek *et al.*, 1988; McGrath and Varshavsky, 1989).

The *ltpgpA* gene is transcribed at low levels in wild type cells, but a 5.5 kb RNA is readily visible in the highly resistant MTX cell lines exhibiting DNA amplification (Figure 5). The abundance of this RNA correlates well with the level of DNA amplification. However, some mutants resistant to high MTX concentration, exemplified here by TarII 3.1000, have neither H circle amplification (or any other detectable form of DNA amplification) nor do they overexpress the *ltpgpA* RNA.

P-glycoprotein genes in kinetoplastida

All P-glycoprotein genes described thus far (with the possible exception of the STE6 yeast gene) are part of a multigene family with at least two members (Endicott and Ling, 1989; van der Bliek and Borst, 1989; Wilson *et al.*, 1989). With a nbs probe (nbsA in Figure 2), we have tested whether more than one P-glycoprotein gene is present in *L.tarentolae*. Lane 1 of Figure 6 shows that more than one restriction



B.

```
Hisp (1)      MMSENKLHVI DLHKRYGGHE VLKGVSLQAR AGDVISIIGS SGSGKSTFLR CINFLEKPSE
nSc  (354)    TSDLTFANVS FSYPSRPSEA VLKNVSLNFS AGQFTFIVGK SGSGKSTLSN LLLRFYDGYN
cSc  (1049)   KPIVSIQNLT FAYPSAPTAF VYKNMNFDMF CGQTLGIIGE SGTGKSTLVL LLTKLYNCEV
nPf  (375)    NKKIEFKNVR FHYDTRKDVE IYKDLSFTLK EGKTYAFVGE SGCGKSTILK LIERLYDPTE
cPf  (1123)   KGKVDIKDVN FRYISRPNVP IYKNLSFTCD SKKTTAIVGE TGSGKSTFMN LLLRFYDLKN
nHul (389)    KGNLEFRNVH FSYPSRKEVK ILKGLNLKVQ SGQTVALVGN SGCGKSTTVQ LMQRLYDPTE
cHul (1032)   EGNVTFGEVV FNYPTRPDIP VLQGLSLEVK KGQTLALVGS SGCGKSTVVQ LLERFYDPLA
nLt  (629)    DEKGEVEEGD REYYQLVSKE LLRNVSLTIP KGKLTMVIGS TGSGKSTLLG ALMGEYSVES
cLt  (1282)   QAGSLVLEGV QMRYREGLPL VLRGVSFQIA PREKVGIVGR TGSGKSTLLL TFMRMVEVCG
                         ▲ ▲        ■ ■ ■▲      ▲    ■■*  ▲* **** ■  ▲■ ▲ ▲

Hisp (69)     GAIIV.NGQN INLVRDKDGQ LKVADKNQLR LLRTRLTMVF QHFNLWS... ..........
nSc  (414)    GSISI.NGHN IQTIDQKLLI ENITVVEQRC TLFNDTLRKN ILLGSTD... ..........
cSc  (1109)   GKIKI.DGTD VNDWNLTSLR KEISVVEQKP LLFNGTIRDN LTYGLQD... ..........
nPf  (435)    GDIIVNDSHN LKDINLKWWR SKIGVVSQDP LLFSNSIKNN IKYSLYSLKD LEAMENYYEE
cPf  (1183)   DHIILKN..D MTNFQDYQNN NNNSLVLKNV NEFSNQSGSA EDYTVFNNNG EILLDDINIC
nHul (449)    GMVSV.DGQD IRTINVRFLR EIIGVVSQEP VLFATTIAEN IRYGR..... ..........
cHul (1092)   GKVLL.DGKE IKRLNVQWLR AHLGIVSQEP ILFDCSIAEN IAYGDNS... ..........
nLt  (689)    GELWA..... .........E RSIAYVPQQA WIMNATLRGN IL........ ..........
cLt  (1342)   GVIHV.NGRE MSAYGLRELR RHFSMIPQDP VLFDGTVRQN VD........ ..........
              ■ ■ ■  ▲   ■               ■ ■    ■▲      ▲

Hisp (107)    .......... .......... ....HMTVLE NVMEAPIQVL GLSKHDARER ALKYLAKVGI
nSc  (460)    .......... .......... .....SVRNA DCSTNENRHL IKDACQMALL DRFILDLPDG
cSc  (1155)   .......... .......... .......... ....EILEIE MYDALKYVGI HDFVISSPQG
nPf  (495)    NTNDTYENKN FSLISNS... MTSNELLEMK KEYQTIKDSD VVDVSKKVLI HDFVSSLPDK
cPf  (1181)   DYNLRDLRNL FSIVSQEPML FNMSIYENIK FGREDATLED VKRVSKFAAI DEFIESLPNK
nHul (493)    .......... .......... .......... ...ENVTMDE IEKAVKEANA YDFIMKLPHK
cHul (1138)   .......... .......... .......... ...RVVSQEE IVRAAKEANI HAFIESLPNK
nLt  (717)    .......... .......... .......... .FFDEERAED LQDVIRCCQL EADLAQFCGG
cLt  (1383)   .......... .......... .......... .PFLEASSAE VWAALELVGL RERVASESEG
                                                          ■            ▲■   ▲

Hisp (143)    DERAQGKYPV HLSGGQQQRV SIARALA.ME PDVLLFDEPT SALDPELVGE VLRIMQQ..L
nSc  (495)    LETLIGTGGV TLSGGQQQRV AIARAFI.RD TPILFLDEAV SALDIVH.RN LLMKAIR..H
cSc  (1181)   LDTRI..DTT LLSGGQAQRL CIARALL.RK SKILILDECT SALDSVS.SS IINEIVK..K
nPf  (552)    YDTLVGSNAS KLSGGQKQRI SIARAIM.RN PKILILDEAT SSLDNKS.EY LVQKTINNLK
cPf  (1301)   YDTNVGPYGK SLSGGQKQRI AIARALL.RE PKILLLDEAT SSLDSNS.EK LIEKTIVDIK
nHul (520)    FDTLVGERGA QLSGGQKQRI AIARALV.RN PKILLLDEAT SALDTES.EA VVQVALD..K
cHul (1165)   YSTKVGDKGT QLSGGQKQRI AIARALV.RQ PHILLLDEAT SALDTES.EK VVQEALD..K
nLt  (746)    LDTEIGEMGV NLSGGQKARV SLARAVY.AN RDVYLLDDPL SALDAHVGQR IVQDVIL..G
cLt  (1412)   IDSRVLEGGS NYSVGQRQLM CMARALLKRG SGFILMDEAT ANIDPAL.DR QIQATVM..S
               ■  ■▲      ■*■** ■■■  ■***■  ▲    ■■■■*■▲▲ ■▲■*       ■■   ■

Hisp (200)    AEEGKTMVVV THEMGFARHV SSHVIFLH
nSc  (551)    WRKGKTTIIL THELSQIESD DYLYLMKE
cSc  (1235)   GPPALLTMVI THSEQMMRSC NSIAVLKD
nPf  (610)    GNENRITIII AHRLSTIRYA NTIFVLSN
cPf  (1359)   DKADKTIITI AHRIASIKRS DKIVVFNN
nHul (576)    ARKGRTTIVI AHRLSTVRNA DVIAGFDD
cHul (1221)   AREGRTCIVI AHRLSTIQNA DLIVVFQN
nLt  (803)    RLRGKTRVLA THQIHLLPLA DYIVVLQH
cLt  (1469)   AFSAYTVITI AHRLHTVAQY DKIIVMDH
                   ▲ ■■■  * ■       ▲ ■ ■
```

Fig. 4. Comparison of the *Leishmania* P-glycoprotein A with other P-glycoproteins. (A) Hydropathy profiles generated by using the algorithm of Kyte and Doolittle (1982) for a window size of 14 residues. The complete *ltpgpA* profile and a scale up of it, for better comparison with the human mdr1 P-glycoprotein, are shown. Hydrophobic and hydrophilic regions fall above and below the centre line respectively. The schematic representations of ltpgpA and mdr1 are shown below the profiles. The transmembrane domains are indicated by hatched boxes and the nbs consensus sequences in solid boxes. (B) Alignment of P-glycoprotein nbs. For comparison the prototype nbs consensus sequence of the bacterial transporter histidine permease (hisp) (Higgins *et al.*, 1986) is also shown. The duplicated nbs of the P-glycoproteins of the yeast *S.cerevisiae* STE6 (nSc and cSc) (McGrath and Varshavsky, 1989), of *P.falciparum* (nPf and cPf) (Foote *et al.*, 1989), of the human mdr1 (nHul and cHul) (Chen *et al.*, 1986) and of *L.tarentolae* (nLt and cLt) are shown. Numbers in parentheses at the left correspond to the position of the first amino acid of the sequence. Gaps were introduced to maximize homology. Asterisks indicate identities among the nine nbs sequences; squares indicate that at least eight of the nine nbs share identical or highly conserved amino acids; and a triangle indicates that at least six of the nine nbs motifs have identical amino acids.

Table I. Transporter proteins containing the nucleotide binding site consensus motif[a]

Protein	Nonduplicated		Protein	Duplicated	
	Host	Substrate		Host	Substrate
brown	*Drosophila*	pteridine[b]	CFRT	human	?
btuD	*Escherichia coli*	vitamin B-12	itpgpA	*Leishmania*	?
chlD	*E.coli*	molybdene	mdr1	human	drugs[c]
fhuC	*E.coli*	iron-ferrichrome	mdr3	human	?
hisp	*Salmonella*	histidine			
hlyB	*E.coli*	hemolysin A	mdr1[d]	mouse	drugs[c]
malK	*E.coli*	maltose	mdr2	and	?
mbpX	Liverwort	?	mdr3	hamster	drugs[c]
nodI	*Rhizobium*	?			
oppD	*Salmonella*	oligopeptide	pfmdr1	*Plasmodium*	drugs[c]
pstB	*E.coli*	phosphate	rbsA	*E.coli*	ribose
white	*Drosophila*	pteridine[b]	STE6	*Saccharomyces*	a-factor

[a] The alignment of the primary sequence of most of the transporter proteins is presented by Riordan *et al.* (1989). Not included were *brown, chlD* and *fhuC* that can be found in Dreesen *et al.* (1988), Johann and Hinton (1987) and Coulton *et al.* (1987) respectively.

[b] Suggested substrate, for more detail see Dreesen *et al.* (1988).

[c] Are able to transport drugs but also probably a yet unidentified physiological substrate. Drug transport by pfmdr1 is likely but not proven.

[d] Nomenclature of Gros *et al.* (1988) for the mouse P-glycoproteins. The mouse *mdr1, 2* and *3* are the equivalent of the hamster *pgp2, 3* and *1* respectively.

fragment in *L.tarentolae* DNA hybridizes with a nbs probe. This hybridization signal resists stringent washes at 65°C, 0.1×SSC (not shown). From further restriction digests and cloning experiments, we infer that at least three P-glycoprotein genes are probably present in *Leishmania* (M.Ouellette, E.Hettema and P.Borst, unpublished observations), but only one of those genes is present on the H circle. The other lanes of Figure 6 show that all the tested representatives of the kinetoplastida contain DNA sequences homologous to the nbs region of *ltpgpA*; in most parasites more than one DNA restriction fragment hybridizes with the probe. This result substantiates the ubiquity of P-glycoprotein genes in eukaryotes.

H circle amplification does not result in multidrug resistance or arsenite resistance

Although the relation between H region amplification and MTX resistance is now well established in *Leishmania* (Hightower *et al.*, 1988; Petrillo-Peioxoto and Beverely, 1988; White *et al.*, 1988), the involvement of the H circle in multidrug resistance (Ellenberger and Beverley, 1989) or arsenite resistance (Katakura and Chang, 1989) is less clear. We have therefore tested several *L.tarentolae* mutants, selected for MTX resistance, for cross-resistance to a range of other drugs. The strains tested included two new mutants of strain TarII, which is devoid of H circles. In one of these, TarII 1.1000, H circles were generated during MTX selection; in the other mutant, TarII 3.1000, a high level of resistance was obtained without H circle formation or over-expression of the chromosomal copy of H circle sequences. The results of the cross-resistance experiments, summarized in Table II, show that the resistant mutants are neither cross-resistant to the lipophilic MTX analogue trimetrexate, nor to arsenite, nor to drugs that are part of the mammalian or *Plasmodium* multidrug resistance spectrum. MTX resistance was also not affected by verapamil, a drug that can reverse multidrug resistance (not shown).

Ellenberger and Beverley (1989) have recently reported that *L.major* strains selected for primaquine, terbinafine or MTX resistance, are (cross-) resistant to primaquine, if the H circle is amplified. Resistance was only 1.2- to 2.0-fold and not obviously related to H circle copy number. We also see slight variations in the sensitivity to primaquine among our mutants (Table II). These variations are clearly not related to the level of *ltpgpA* transcripts (Figure 5) and their significance is doubtful.

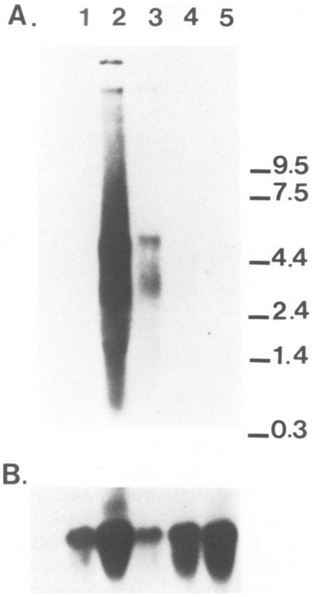


Fig. 5. Expression of the *L.tarentolae* P-glycoprotein gene A in wild type and mutant strains. (**A**) Total RNA of wild type and mutant strains was isolated, electrophoresed on a 1% agarose gel, blotted and hybridized with the probe nbsA (see Figure 2). Hybridization in the high mol. wt region is probably due to free H circle DNA that contaminates the RNA. (**B**) The gel was rehybridized with a tubulin probe to monitor the amounts of RNA layered on the gel. The mol. wt marker is the RNA ladder from BRL. 1, TarII 3.1000; 2, TarVIa 1000; 3, TarII 1.1000; 4, TarII 1.200; 5, TarII WT.

Discussion

Our results show that the H circle of *L.tarentolae* contains a gene that encodes a protein with the hallmarks of a P-glycoprotein. Sequences of P-glycoproteins from mammals (Chen *et al.*, 1986; Gerlach *et al.*, 1986; Gros *et al.*, 1986, 1988; van der Bliek *et al.*, 1988). *Drosophila* (J.Croop, personal communication), *Caenorhabditis* (C.Lincke and P.Borst, unpublished data), *Plasmodium* (Foote *et al.*, 1989; Wilson *et al.* 1989) and yeast (McGrath and Varshavsky, 1989) are now available and among these the sequence of the ltpgpA is the most divergent. It is therefore not surprising that the gene product does not seem to be able to confer resistance to large hydrophobic drugs, like some of the mammalian P-glycoproteins and the *Plasmodium* P-glycoprotein. Amplification of H circle like plasmids in *L.major* has been linked to resistance to terbinafine and primaquine (Ellenberger and Beverley, 1989) and in *L.mexicana amazonensis* to arsenite (Katakura and Chang, 1989). As there are probably at least three P-glycoprotein genes in *Leishmania*, it is possible that these H circles differ from the ones studied here, but H circle amplification in these mutants may also be unrelated to the drug resistance observed.

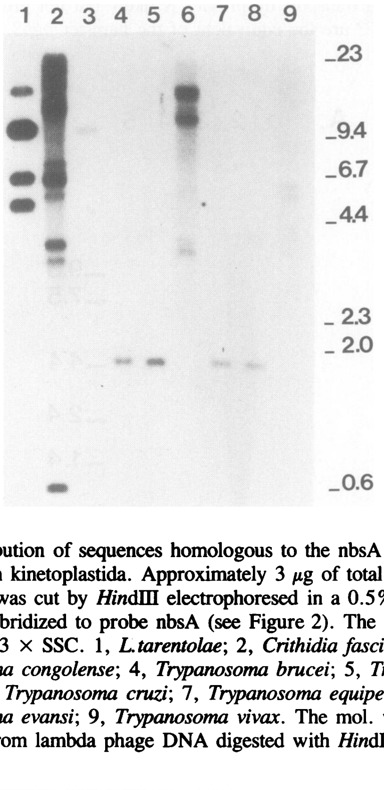


Fig. 6. Distribution of sequences homologous to the nbsA sequence of *L.tarentolae* in kinetoplastida. Approximately 3 μg of total DNA of each parasite was cut by *Hind*III electrophoresed in a 0.5% gel, blotted and hybridized to probe nbsA (see Figure 2). The final wash was at 65°C, 3 × SSC. 1, *L.tarentolae*; 2, *Crithidia fasciculata*; 3, *Trypanosoma congolense*; 4, *Trypanosoma brucei*; 5, *Trypanosoma gambiense*; 6, *Trypanosoma cruzi*; 7, *Trypanosoma equiperdum*; 8, *Trypanosoma evansi*; 9, *Trypanosoma vivax*. The mol. wt marker was derived from lambda phage DNA digested with *Hind*III.

There are now several observations linking H circle amplification to MTX resistance. Amplification has been induced in several different *Leishmania* strains by growth in MTX (Beverley *et al.*, 1984; Hightower *et al.*, 1988; Petrillo-Peioxoto and Beverley, 1988; White *et al.*, 1988). The amplification is stably maintained after full adaptation to growth in MTX and slowly decreases again when MTX is eliminated from the growth medium. This argues against the possibility that H circle amplification is a generalized stress response. The fact that highly resistant mutants can be obtained that do not contain H circles (Table I), is another argument against this possibility. Our attempts to prove directly that H circles confer MTX resistance, by electroporation of these circles into *L.tarentolae*, have failed thus far, however.

How a P-glycoprotein might contribute to MTX resistance remains unclear. It is conceivable that ltpgpA extrudes MTX from the cell, as some of the proteins in Table I transport hydrophilic compounds. The *brown* and *white* proteins of *Drosophila* are even thought to transport pteridines (Dressen *et al.*, 1988), structurally related to MTX. It is even possible that a P-glycoprotein would be discriminating enough to remove MTX from the cell without also depleting it of reduced folates, as folates are polyglutamylated (Santi *et al.*, 1987), whereas MTX is not in *Leishmania* (Ellenberger *et al.*, 1989). However if a P-glycoprotein were involved in MTX resistance one would expect an increased efflux of the drug. Transport studies (Ellenberger and Beverley, 1987, 1989) indicate that there is no clear relation between H circle amplification and MTX efflux in *L.major*. It is therefore not clear whether there is a link between *ltpgpA* and MTX resistance and another gene on the H circle may therefore contribute to MTX resistance.

We have previously proposed that the generation of H circles with their unusual inverted duplication is not a laboratory artefact, but part of a physiological response system, allowing *Leishmania* to produce additional H region copies as need arises (White *et al.*, 1988). The H region might contain genes required in the defense against foreign toxic compounds, as often found on plasmids of other microorganisms. Our results give some credibility to this hypothesis. We have found that authentic H circles can be generated *de novo* (Table I) and we have verified by

Table II. Relative drug resistance and H circles in *L.tarentolae*[a]

Strains	H circle[b]	Drugs[c]							
		MTX	TMQ	VCR	ADR	CLQ	AS	PRQ	PUR
VIa WT	+	1	ND[e]	ND	1	1	ND	1	1
VIA MTX1000	+++	40	ND	ND	0.75	1.2	ND	2	1
II WT	−	0.5[d]	1	1	1	1	1	1	1
II 1.1000	++	>40	0.5	1	0.5	0.3	1.2	1	1
II 3.1000	−	>40	0.5	1	0.5	0.7	1.1	1.8	1

[a] The relative drug resistance values were obtained by dividing the 50% growth inhibition value of the strain tested by the 50% growth inhibition value of the wild type cells. A value of 1 indicates that the mutant and the wild type have the same sensitivity to the drug.

[b] The level of DNA amplification was determined by DNA hybridization with a H circle probe. The copy number of free H circle was undetectable (−) or ~100 copies of the circle (+++) per cell.

[c] Abreviations for drugs are MTX, methotrexate; TMQ, trimetrexate; VCR, vincristine; ADR, adriamycin (doxorubicin); CLQ, chloroquine; AS, arsenite; PRQ, primaquine; PUR, puromycin.

[d] For MTX, the relative drug resistance of TarII WT was compared with strain TarVIa WT.

[e] Not determined.

sequence analysis that the chromosomal H region is flanked on one side by the inverted duplication required for generating H circles according to the 'panhandle' scenario of White *et al.* (1988) (unpublished results). The presence of a P-glycoprotein gene in the H circle is in line with a possible function of the H region in defense. It will be of interest to determine what other genes are present in this region and how the region is copied from the genome.

Materials and methods

Cell lines and culture

The two parental cell lines, TarII and TarVIa, were obtained from the American Type Culture Collection (ATCC No. 30267) and from Dr F.R. Opperdoes (ICP, Brussels) respectively (White *et al.*, 1988). Cells were grown in SDM-79 medium as described (Hoeijmakers *et al.*, 1981) in the presence or absence of MTX (Emthexate, Pharmachemie B.V., Haarlem, Holland). MT resistant mutant TarVIa MTX1000 was described by White *et al.* (1988). The TarII 1.1000 and TarII 3.1000 are two independent mutants derived from TarII. Cells were adapted in 3–6 months to 1000 μM MTX by passaging them in increasing drug concentrations at steps of 50, 200, 500 and 1000 μM MTX. We have previously reported that the TarII wild type strain is devoid of H circles and we infer that the circles present in TarII 1.1000 have arisen *de novo*. This is supported by two additional observations: (i) additional mutants containing H-circles were selected with MTX from cloned TarII populations without circles; and (ii) the exact positions of the borders of the inverted repeats differ in the H circles from different MTX resistant mutants derived from cloned TarII populations. Growth curves were constructed by measuring absorbance at 600 nm. Relative drug resistance was measured by dividing the 50% growth inhibition value of the strain tested by the 50% growth inhibition value of the wild type cell. The drugs used were obtained from the following suppliers: adriamycin, chloroquine, primaquine and vincristine (Sigma); sodium arsenite (Merck); puromycin (Boehringer); trimetrexate was the generous gift of Dr G.Jansen (Academic Hospital, Utrecht).

Nucleic acid isolation and blotting

Hirt supernatants (Hirt, 1967) were prepared as described previously (White *et al.*, 1988). Total DNA was isolated as described (Bernards *et al.*, 1981). Total parasite DNA was digested by *Hin*dIII, electrophoresed through an agarose gel and blotted onto nitrocellulose. The blot was hybridized with probe nbsA (see Figure 2) labelled by random priming (Feinberg and Vogelstein, 1983). Total RNA was prepared by cell lysis in guanidinium isothiocyanate and pelleted through a caesium chloride cushion (MacDonald *et al.*, 1987). Total RNA of *L.tarentolae* was electrophoresed in a 2.2 M formaldehyde containing agarose gel, blotted onto nitrocellulose and hybridized in 40% formamide to the labeled probe nbsA (Maniatis *et al.*, 1982). For both the DNA and RNA blot, final post hybridization washes were at 3 × SSC, 65°C (1 × SSC is 0.15 M NaCl and 0.015 M sodium citrate, pH 7.0).

DNA sequence analysis

Double stranded plasmid DNAs, derivatives of pGEM3 or pGEM4 (Promega), were sequenced using the dideoxy method (Sanger *et al.*, 1977). In addition to the SP6 and T7 sequencing primers, oligonucleotides were synthesized using phosphoramidite chemistry on a fully automatic synthesizer (Cyclone DNA synthesizer, Biosearch) and were used as internal primers to complete the sequence on both strands. Computer analysis of the nucleotide sequence was performed using the software package of the University of Wisconsin Genetics Computer Group (UWGCG) (Devereux *et al.*, 1984). This sequence will appear in the EMBL/GenBank/DDBJ nucleotide sequence databases under the accession number X17154.

Acknowledgements

We thank Dr Fred Opperdoes (ICP, Brussels) for the *L.tarentolae* strain TarVIa, Dr G.Jansen (Academic Hospital, Utrecht) for trimetrexate, Dr Ted White for several of the plasmids used in this study and Dr Carsten Lincke for the human *mdr1* cDNA probe. We also thank our colleagues of both the 'tryps' and 'mdr' group of the Borst laboratory for interest and advice throughout the course of this work and for critical reading of the manuscript. This work was supported in part by a grant from the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (NWO) and by a NATO (Canada) post-doctoral fellowship to M.O.

References

Bernards,A., Van der Ploeg,L.H.T., Frasch,A.C.C., Borst,P., Boothroyd,J.C., Coleman,S. and Cross,G.A.M. (1981) *Cell*, **27**, 497–505.

Beverley,S.M., Coderre,J.A., Santi,D.V. and Schimke,R.T. (1984) *Cell*, **38**, 431–439.

Chen,C., Chin,J.E,. Ueda,K., Clark,D.P., Pastan,I., Gottesman,M.M. and Roninson,I.B. (1986) *Cell*, **47**, 381–389.

Coulton,J.W., Mason,P. and Allatt,D.D. (1987) *J. Bacteriol.*, **169**, 3844–3849.

Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387–395.

Dreesen,T.D., Johnson,D.H. and Henikoff,S. (1988) *Mol. Cell. Biol.*, **8**, 5206–5215.

Ellenberger,T.E. and Beverley,S.M. (1987) *J. Biol. Chem.*, **262**, 13501–13506.

Ellenberger,T.E. and Beverley,S.M. (1989) *J. Biol. Chem.*, **264**, 15094–15103.

Ellenberger,T.E., Wright,J.E., Rosowsky,A. and Beverley,S.M. (1989) *J. Biol. Chem.*, **264**, 15960–15966.

Endicott,J.A. and Ling,V. (1989) *Annu. Rev. Biochem.*, **58**, 137–171.

Feinberg,A.P. and Vogelstein,B. (1983) *Anal. Biochem.*, **132**, 6–13.

Friedlander,M. and Blobel,G. (1985) *Nature*, **318**, 338–343.

Foote,S.J., Thompson,J.K., Cowman,A.F. and Kemp,D.J. (1989) *Cell*, **57**, 921–930.

Gerlach,J.H., Endicott,J.A., Juranka,P.F., Henderson,G., Sarangi,F., Deuchars,K.L. and Ling,V. (1986) *Nature*, **324**, 485–489.

Gottesman,M.M. and Pastan,I. (1988) *J. Biol. Chem.*, **263**, 12163–12166.

Gros,P., Croop,J. and Housman,D. (1986) *Cell*, **47**, 371–380.

Gros,P., Raymond,M., Bell,J. and Housman,D. (1988) *Mol. Cell. Biol.*, **8**, 2770–2778.

Higgins,C.F., Hiles,I.D., Salmond,G.P., Gill,D.R., Downie,J.A., Evans,I.J., Holland,I.B., Gray,L., Buckel,S.D., Bell,A.W. and Hermodson,M.A. (1986) *Nature*, **323**, 448–450.

Hightower,R.C., Ruiz-Perez,L.M., Lie Wong,M. and Santi,D.V. (1988) *J. Biol. Chem.*, **263**, 16970–16976.

Hirt,B. (1967) *J. Mol. Biol.*, **26**, 365–369.

Hoeijmakers,J.H.J., Snijders,A., Janssen,J.W.G. and Borst,P. (1981) *Plasmid*, **5**, 329–350.

Johann,S. and Hinton,S.M. (1987) *J. Bacteriol.*, **169**, 1911–1916.

Katakura,K. and Chang,K.-P. (1989) *Mol. Biochem. Parasitol.*, **34**, 189–192.

Kaur,K,. Coons,T., Emmett,K. and Ullman,B. (1988) *J. Biol. Chem.*, **263**, 7020–7028.

Kozak,M. (1987) *Nucleic Acids Res.*, **15**, 8125–8148.

Kyte,J. and Dootlittle,R.F. (1982) *J. Mol. Biol.*, **157**, 105–132.

MacDonald,R.J., Swift,G.H., Przybyla,A.E. and Chirgwin,J.M. (1987) *Methods Enzymol.*, **152**, 219–227.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

McGrath,J.P. and Varshavsky,A. (1989) *Nature*, **340**, 400–404.

Petrillo-Peioxoto,M.L. and Bevereley,S.M. (1988) *Mol. Cell. Biol.*, **8**, 5188–5199.

Riordan,J.R., Rommens,J.M., Kerem,B.-S., Alon,N., Rozmahel,R., Grzelczac,Z., Zielenski,J., Lok,S., Plavsic,N., Chou,J.-L., Drumm, M.L., Iannuzzi,M.C., Collins,F.S. and Tsui,L.-C. (1989) *Science*, **245**, 1066–1073.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.

Santi,D.V., Nolan,P. and Shane,B. (1987) *Biochem. Biophys. Res. Commun.*, **146**, 1089–1092.

van der Bliek,A.M. and Borst,P. (1989) *Adv. Cancer Res.*, **52**, 165–203.

van der Bliek,A.M., Kooiman,P.M., Schneider,C. and Borst,P. (1988) *Gene*, **71**, 401–411.

Von Heine,G. (1985) *J. Mol. Biol.*, **184**, 99–105.

White,T.C., Fase-Fowler,F., van Luenen,H., Calafat,J. and Borst,P. (1988) *J. Biol. Chem.*, **263**, 16977–16983.

Wilson,C.M., Serrano,A.E., Wasley,A., Bogenschutz,M.P., Shankar,A.H. and Wirth,D.F. (1989) *Science*, **244**, 1184–1186.

Received November 14, 1989.