Sequence coding for a sexual stage specific protein of *Plasmodium falciparum*

Marian C.Bruce, David A.Baker¹, Pietro Alano, Neil C.Rogers¹, Patricia M.Graves², Geoffrey A.T.Targett¹ and Richard Carter

Department of Genetics, University of Edinburgh, West Mains Road, Edinburgh EH9 3JN,
¹Department of Medical Parasitology, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK and ²Queensland Institute of Medical Research, Bramston Terrace, Herston, Brisbane, Queensland 4006, Australia

Submitted May 21, 1990

EMBL accession no. X53030

The human malarial parasite *Plasmodium falciparum* has two developmentally different intra-erythrocytic forms: asexual parasites which are responsible for the vast asexual multiplication in the human host and the sexual forms, gametocytes, which are the precursors of male and female gametes which form in the midgut of the mosquito vector when gametocytes are ingested during a blood meal. We report here the cloning of a sequence encoding a sexual stage-specific protein. Two monoclonal antibodies (Mabs), 2G7 and 93A3A2, react in immunofluorescence assays, with gametocytes but not with asexual blood stage parasites of *P. falciparum*. Both Mabs react with a protein migrating between Mr 17000 and 18000 on Western blot of *P. falciparum* gametocyte protein. These Mabs were used to isolate two overlapping fragments of DNA.

A gametocyte specific cDNA library in lambda gt11 was screened with one monoclonal antibody 2G7, and 19 positive clones were isolated from 2×10^4 plaques. The clone containing the largest, 0.4 kb insert was sequenced revealing a fragment containing a long open reading frame and a termination codon,

followed by a highly AT-rich region. This fragment forms the 3' end of the sequence shown below. A second, independently generated monoclonal antibody, 93A3A2, was used to screen RsaI and AluI genomic expression libraries in lambda gt11. From 4×10^4 plaques, two positives were isolated. One contained a single open reading frame of 222 bp which forms the 5' end of the sequence below. These two clones share an overlapping region of 57 bp, which is underlined in the combined sequence below. The deduced amino acid sequence gives a molecular weight of 15.4 kD, which is close to the molecular weight of the protein estimated on Western blot. The boxed areas marked on the amino acid sequence refer to hydrophobic regions, the first of which could represent a signal peptide. The methionine at amino acid position 3, therefore, may be the translation start codon. Evidence for this sequence being transcribed in vivo comes from Northern blot analysis. An oligonucleotide, designed from the sequence reported here, hybridises to a single 1.3 kD band on *P. falciparum* gametocyte mRNA.

```

GCTTTAATGCTTATATTCTTCGCTTTTGCAAACCTGGTATTATCAGATGCAAATGACAAA   60
A L M L I F F A F A N L V L S D A N D K

GCAAAAAAGCCGCTGGAAAAGGATCCCTTCAACTTTGCAAACCCAGGAAGTTCTTCA   120
A K K P A G K G S P S T L Q T P G S S S

GGTGCCTCTCTTCATGCTGTTGGACCTAATCAAGGTGGACTATCTCAAGGTCTTTCTGGA   180
G A S L H A V G P N Q G G L S Q G L S G

AAAGATTCTGCTGACAAAATGCCTTTAGAAACTCAGCTAGCTATAGAAGAAATCAAGAGC   240
K D S A D K M P L E T Q L A I E E I K S

TTATCCAATATGTTAGATAAAAAACGACAGTTAACAGAACTTAATCATAAGTACTGCT   300
L S N M L D K K T T V N R N L I I S T A

GTCACAAATATGATCATGTTGATCATATTATCTGGTATAGTTGGATTTAAAGTTAAAAAA   360
V T N M I M L I I L S G I V G F K V K K

ACGAAGAACGCAGATGATGATAAAGGAGATAAGGATAAGGACAAGGATAATACAGATGAA   420
T K N A D D D K G D K D K D K D N T D E

GGAGACGAAGGAGATGATTCTTAAATGTATATATATATATATATATATATATT   473
G D E G D D S

```