Molecular and Biochemical Parasitology, 39 (1990) 47–60
Elsevier

MOLBIO 01274

Characterisation of the gene encoding a 104-kilodalton microneme-rhoptry protein of *Theileria parva*

Keith Paul Iams, John R. Young, Vishvanath Nene, Jagruti Desai*, Paul Webster, Onesmo K. ole-MoiYoi and Anthony J. Musoke

International Laboratory for Research on Animal Diseases, Nairobi, Kenya

(Received 3 July 1989; accepted 25 September 1989)

A neutralizing antiserum, C16, raised against sporozoites of *Theileria parva parva* was used to screen a λgt11 expression library of *T. parva parva* (Muguga) genomic DNA fragments. Proteins encoded by one phage clone, λgTpS-17, were reactive with the C16 antiserum. Detailed characterisation of the DNA insert showed it to encode determinants found on four theilerial antigens of approximately 104, 90, 85 and 35 kDa. The sequence encoded by the clone is expressed during sporogony as a single RNA transcript of about 3000 nucleotides. On sequencing a portion of the 5000-bp insert, an open reading frame of 2772 bp was revealed that encoded a 104-kDa protein. Immunoscreening a library of subfragments of the DNA insert with the original antiserum localised sequences encoding the dominant antigenic determinants to an 800-bp stretch of DNA at the 3′ end of the open reading frame. Sequence data from three subclones spanning this region show portions of the antigenic domains to be unusually rich in proline residues which are repeated every three amino acids. These repeats often take the form X-S(T)-P or X-K(R)-P. Antibodies directed against each of the three subclones recognize the 104- and 35-kDa antigens and different combinations of the 90- and 85-kDa antigens, suggesting that the smaller proteins are derived from the 104-kDa antigen by limited proteolysis occurring at the carboxyl terminus end of the protein. In immunoelectron micrographs the antigen is associated with the microneme/rhoptry complexes of the sporozoite.

Key words: *Theileria parva*; East Coast fever; Parasite antigen; Sporozoite; Genomic expression library, λgt11; Immunoscreening; Epitope mapping

Introduction

East Coast fever, a disease complex of cattle caused by the hemoprotozoan parasite *Theileria parva*, presents a major obstacle to livestock development in East and Central Africa. The disease is transmitted by the brown ear tick *Rhipicephalus appendiculatus*. The mammalian-infective stage of *Theileria*, the sporozoite, enters cells by employing a ligand-receptor endocytotic process [1] which has been shown to require little energy. Even dead sporozoites are capable of entering lymphocytes [2]. Subsequent stable establishment in the host cell is a complex process that has been shown to involve several parasite structures including the rhoptries and micronemes. At the ultrastructural level, these are visualized as electron-dense organelles which can be differentiated by their size and shape; micronemes are intermediate sized, membrane-bound vesicles, whereas rhoptries are large, pear-shaped structures.

In *Theileria*, diminution of the electron-dense contents of the rhoptries has been observed concurrently with the dissolution of the host cell membrane which encapsulates the endocytosed parasite. Thus, the parasite comes to lie free in

Correspondence address: Keith P. Iams, ILRAD, P.O. Box 30709, Nairobi, Kenya.

*Deceased

Note: Nucleotide sequence data reported in this paper have been submitted to the GenBank™ Data Bank with the accession number M29954

Abbreviations: Tris, 2-amino-2-hydroxymethylpropane-1,3-diol; IPTG, isopropyl β-D-thiogalactoside.

0166-6851/90/$03.50 © 1990 Elsevier Science Publishers B.V. (Biomedical Division)

the cytoplasm, and cannot be destroyed by interaction with host lysosomes. In this process, the precise role of the micronemes has not yet been defined, although in a related sporozoan, *Plasmodium* [3], both micronemes and rhoptries are considered to play a role in promoting successful entry of the parasite into host cells.

In other sporozoans, such as *Plasmodium* and *Toxoplasma*, proteins located within the rhoptries have been found to be highly immunogenic. In *Plasmodium falciparum*, a number of different antigens have been described including proteins of approximately 225 kDa [4], 150/130 kDa [5], 140/130/105 kDa [6–8] and 80/40 kDa [9,10]. In *Plasmodium yoelii* [11,12] a 235-kDa antigen has been described, while in *Toxoplasma gondii* [13], 60/55-kDa rhoptry antigens have been identified. Additionally, the 235-kDa *P. yoelii* [11,12] the 80-kDa *P. falciparum* [14] and an uncharacterized *T. gondii* rhoptry protein [15] have been observed to elicit protective antibody responses which are thought to function by thwarting invasion of the host cell by the parasite.

In this paper we describe the molecular characterisation of a strongly antigenic protein which is localised in the microneme/rhoptry complexes of the sporozoite stage of *T. parva*.

Materials and Methods

Parasite stabilates. Calves, 6 to 12 months of age, were infected by inoculation with a sporozoite stabilate prepared as previously described [16]. A *T. parva parva* (Muguga) stabilate ILRAD836 was used as source material for all molecular and immunological procedures described.

Bacterial strains and vectors. Escherichia coli strains Y1089 and Y1090 r^- m^+ and dephosphorylated λgt11 arms were obtained from Promega Biotec (Madison, WI, U.S.A.).

Isolation of RNA from normal and Theileria-infected tick salivary glands. Salivary glands were dissected from ticks fed for 4 days, flash-frozen and stored in liquid nitrogen. RNA was extracted using the hot phenol-SDS procedure [17].

DNA and RNA blotting. DNA and RNA were transferred onto nitrocellulose filters by methods previously described [18]. DNA was radiolabelled by nick translation using [α-^{32}P]dATP (>3000 Ci $mmol^{-1}$; Amersham, Amersham, U.K.) as described previously [19].

Subcloning of the pGT-17 DNA insert in λgt11. Following *Eco*RI digestion, the 5-kb DNA insert in pgT-17 was separated from the plasmid vector by agarose gel electrophoresis and purified using a Geneclean kit (Bio-101, La Jolla, CA). The DNA insert was sonicated on ice for 5 min using a Branson Sonifier (Branson Sonic Power Co., Danbury, CT, U.S.A.) with micro-tip set on maximum wattage. The average size of the resulting fragments was 200 bp. Fragments were ligated to *Eco*RI linkers and inserted into dephosphorylated λgt11 vector arms as previously described [20].

Immunoscreening of λgt11 expression libraries. Immunoscreening was carried out as described by Young et al. [20]. Antigen-antibody complexes were identified by application of a rabbit anti-bovine antibody conjugated to horseradish peroxidase (ICN Immunobiologicals, Lisle, IL, U.S.A.).

Antibody elution from nitrocellulose filters. Antibodies were selected from whole antiserum essentially as described by Ozaki et al. [21]. However, neutralization of 5 ml of 0.2 M glycine following antibody elution was carried out with 800 μl of a mixture containing 350 μl of 2 M Tris, pH 10.2, 500 μl 10 × TS (10 mM Tris-HCl, pH 8.0/150 mM NaCl), 5 μl 10% NaN_3 and 0.25 g of ovalbumin.

SDS-polyacrylamide gels and Western blotting. Polyacrylamide gels were prepared and run as described by Laemmli [22]. 7.5% acrylamide gels were employed for resolution of proteins in salivary gland and bacterial lysates. Western blotting was carried out in a Bio-Rad Trans-blot cell according to the manufacturer's specifications (Bio-Rad, Richmond, CA, U.S.A.). Western blots were incubated in antiserum that was diluted in TS containing 5% ovalbumin. Blots were developed using ^{125}I-labelled protein G (Amersham).

Preparation of lysogens. E. coli strain Y1089 was infected with recombinant λgt11 at a m.o.i. of 10:1. Cells were spread on NZYCM plates containing 100 μg ampicillin ml^{-1}. Colonies appearing after a 24-h incubation period at 37°C were picked onto duplicate plates which were then incubated at 30°C or 42°C. Lysogens were chosen from those colonies that grew at 30°C but not 42°C. The lysogens were grown overnight at 30°C in 1.0 ml of L-broth containing 100 μg ampicillin ml^{-1}. From an overnight culture, 200 μl overnight culture were added to 0.8 ml of L-broth and grown for a further 2 h at 30°C. Fusion peptide production was initiated by increasing the incubation temperature to 45°C for 20 min, followed by addition of isopropyl β-D-thiogalactoside (IPTG) to a concentration of 10 mM. Growth was continued at 37°C for a further 2.5 h and cells were then harvested by centrifugation at 3000 × *g*. Bacterial pellets were resuspended in 150 μl of TE buffer (10 mM Tris-HCl, pH 8.0/0.1 mM EDTA), solubilized in an equal volume of 2 × sample buffer (0.1 M Tris, pH 6.8/2% SDS/2.0% β-mercaptoethanol, 20% glycerol). Samples were boiled for 3 min at 94°C and subjected to electrophoresis on 7.5% SDS-polyacrylamide gel. β-Galactosidase fusion proteins were made visible by staining the gel with Coomassie Blue 250 or by immunoscreening of Western blots using anti-β-galactosidase antibodies.

Purification of fusion proteins from E. coli lysates. For purification of fusion proteins from bacterial lysates, the bacterial pellets were resuspended in one-tenth volume of buffer containing 20 mM Tris-HCl, pH 7.4/10 mM $MgCl_2$/1.6 M NaCl/10 mM 2-mercaptoethanol. The suspension was freeze-thawed three times and then sonicated on ice for 5 min using a Branson Sonifier micro-tip sonicator set on maximum. The sonicate was cleared by centrifugation at 15000 × *g* for 30 min at 4°C. For purification of the fusion protein, the cleared lysate was applied to a β-galactosidase affinity column prepared and used as described by Ullmann [23].

DNA sequencing. Subcloned DNA fragments were treated with BAL 31 exonuclease (BRL, Middlesex, U.K.), transferred into M13mp18 and sequenced by the Sanger dideoxy chain termination procedure [24]. Each strand was sequenced at least five times and all restriction enzyme sites used for subcloning fragments were sequenced across.

Immunoelectron microscopy of sporozoites. Thin cryosections of partially purified sporozoite preparations were sectioned and labelled as described previously [1].

Results

Southern and Northern blot analyses of phage clone λgTpS-17. Antiserum C16, which was raised by immunising a calf with freeze-thawed lysates of *T. parva parva* sporozoites, and which neutralised sporozoite infectivity in vitro, was used to screen a λgt11 genomic expression library of sheared *T. parva parva* (Muguga) piroplasm DNA. From this screening, several clones were identified that reacted with the antiserum. The peptide product of one clone, λgTpS-17, selected antibodies from whole C16 antiserum that recognised four sporozoite antigens of approximately 104, 90, 85 and 35 kDa. This clone was selected for more detailed characterisation.

The 5000-bp DNA fragment contained in λgTpS-17 was transferred into the plasmid vector pUC-19 and designated pgT-17. Plasmid pgT-17 was radiolabelled by nick translation and employed as a probe to analyse *T. parva parva* (Muguga) DNA restriction fragments on Southern blots. An ethidium bromide-stained agarose gel of *Eco*RI digested purified *T. parva parva* (Muguga) piroplasm DNA is shown in Fig. 1A (lane a). A Southern blot of this gel probed with radiolabelled pgT-17 (lane b) showed hybridization to a 9.3-kb fragment. The probe was next employed on a Northern blot of uninfected and *Theileria*-infected tick salivary gland RNA to determine if the cloned sequence was transcribed during the sporogony of *T. parva* in the tick vector. RNA preparations were made from tick glands isolated on day four following initiation of tick feeding on rabbits. Fig. 1B (lanes a and b) shows ethidium bromide-stained patterns of uninfected and *Theileria*-infected tick salivary gland total RNA, respectively. A duplicate RNA gel

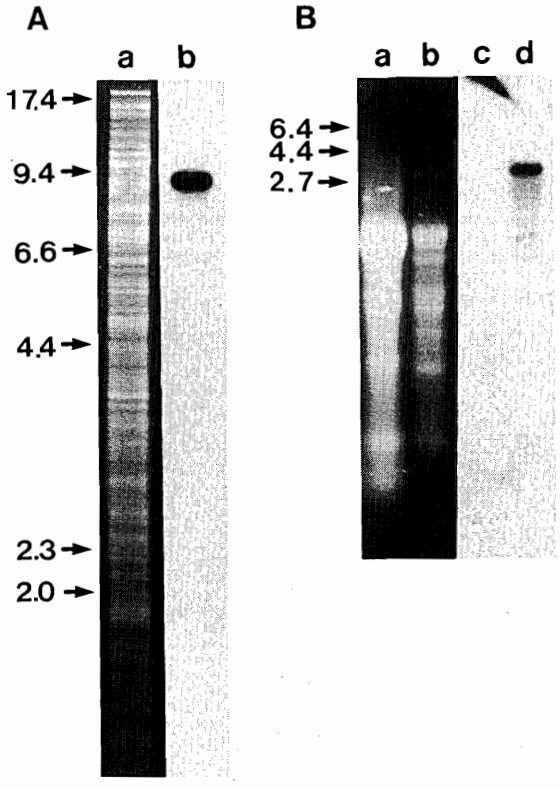

Fig. 1. (A) Ethidium bromide-stained pattern of *Eco*RI-digested *T. parva parva* (Muguga) piroplasm DNA (lane a) which was transferred onto nitrocellulose and probed with nick translated pgT-17 DNA (lane b). Molecular weight markers, in kb are shown on the left of the pattern. (B) Ethidium bromide-stained patterns of uninfected (lane a) and *T. parva parva* (Muguga)-infected tick salivary gland RNA (lane b) were transferred onto nitrocellulose and probed with nick-translated pgT-17 DNA (lanes c and d, respectively). Size markers of single-stranded DNA fragments are shown on the left of the patterns.

was transferred onto nitrocellulose and probed with radiolabelled pgT-17. There was no detectable hybridisation with RNA from uninfected tick salivary glands (lane c). However, a transcript of about 3000 nucleotides was identified in the RNA prepared from the *Theileria*-infected salivary glands (lane d). This transcript is approximately the size expected for an antigen of 100–120 kDa, and confirms that a sequence contained within the recombinant plasmid, pgT-17, is actively transcribed during sporogony.

Sequence of the gene encoded by λgTpS-17. A 4608 bp segment of DNA representing a portion of λgTpS-17 was sequenced. A stop codon map of the six possible reading frames of the sequence, shown in Fig. 2, revealed one open reading frame of sufficient size to encode an antigen of approximately 100 kDa. In Fig. 3, the complete nucleotide sequence of the 4608-bp segment is shown. Starting with the first in-frame ATG initiation codon at base pair 452, the open reading frame extends for 2772 bp and encodes 924 amino acids, having a combined M_r of 104 000. This corresponds well with the estimated size of the largest antigen identified by antibodies selected by λgTpS-17 protein. The initiation codon is followed by a stretch of 15 hydrophobic residues within the first 19 amino acids. The open reading frame terminates at position 3223 with a TAA termination codon preceded by a stretch of 15 hydrophobic residues among the final 19 amino acids. A consensus polyadenylation signal, AATAAA, appears downstream from the presumptive termination signal at position 3306. The entire open reading frame contains 40.8% G+C.

Fig. 2. Stop codon map of the sequenced portion of clone pgT-17. Vertical lines identify stop codons in each of the six possible reading frames. An open reading frame of 2773 bp is identified in frame 2. The scale marks are each 40 bp.

The non-coding region of the fragment has a much lower (30.4%) G+C content.

Localisation and characterisation of the antigenic determinants in λgTpS-17. In order to localise more precisely the antigenic regions encoded by the insert in pgT-17, the 5-kb DNA insert was purified away from the plasmid, sonicated to an average size of 200 bp and re-inserted into the λgt11 expression site. Approximately 2000 clones from this library, containing much smaller DNA inserts, were immunoscreened with the C16 antiserum, and a number of positive clones was isolated. Eight of the subcloned fragments were transferred into M13mp18 and partially sequenced from each end. These sequences are localised on the restriction map of the pgT-17 DNA insert in Fig. 4. Beneath the map is indicated the location of the open reading frame, as well as the locations of the eight subclones. Arrowheads designate the transcriptional direction of the open reading frame encoding the presumptive 104-kDa protein, and the transcriptional orientation of three of the subclones (1, 5 and 9) from the *lacZ* promoter of λgt11. The eight subclones are all located at the 3′ end of the gene and encode three adjacent portions of the antigen. The three portions of the protein express an undefined number of individual determinants, and each will therefore be designated as an antigenic domain. Antigenic domain I is defined by a common sequence in subclones 1, 3, 4, 6 and 8 while antigenic domain III is encoded within subclones 2 and 9. Sequence encoded by subclones 5 spans a portion of domain III and extends toward domain I possibly expressing determinants constituting an additional distinct antigenic domain (II). The extent of the three domains, as delineated by the smallest subcloned fragment, is indicated in Fig. 4 beneath the eight mapped subclones.

Three of the subclones (1, 5 and 9), encoding determinants from all three antigenic domains, were chosen for more detailed characterisation. Delineated within the open reading frame of Fig. 3, are the sequences encoded by subclone 1 (---), 5 (===) and 9 (+++). Analysis of the predicted amino acid sequences encoded by the subclones reveals several unusual characteristics. A portion of both subclone 1 and 9 have an unusually large number of proline residues, which are imprecisely repeated at three amino acid intervals (Fig. 3, ∧∧∧). In subclone 1, the repeats, extending from amino acid residue 691 to residue 729, include 15 prolines, which represent 37% of the sequence. The predicted peptide would also contain 27% of the hydroxyl amino acid residues, serine and threonine. In subclone 9, the repeats extend from amino acid residue 884 to residue 903, and contain 8 proline residues, representing 36% of the sequence. These repeat units are again irregular and, unlike those in subclone 1, contain 48% basic lysine or arginine residues in the first and second amino acid positions. Besides proline, lysine, and arginine, the only other amino acid residue represented in this region is serine. The organisation of the repeat domains found in the two clones is more clearly shown in Fig. 5, where the proline residues have been aligned one above the other.

The repetitiveness of portions of the subclones is not reflected in the DNA sequence. Only one 12-bp sequence occurs three times within subclone 9. This repeat, CCA AAG AAA CCA encodes the three Pro-Lys-Lys-Pro tetrapeptides found in subclone 9, one occurring from nucleotide 3101 to 3112, and a pair of tetrapeptides sharing a common proline residue from bp 3140 to 3160.

Determinants encoded by subclone 1 do not cross-react with those encoded by subclones 5 and 9. Antibodies selected by peptides encoded by each of the three subclones were used to test for the presence of cross-reacting determinants within the three peptide sequences. Selected antibody was applied to Western-blotted fusion peptides made from each of the three subclones. Shown in Fig. 6A is a Coomassie Blue-stained acrylamide gel of lysates made from the induced lysogens of subclones 1, 5 and 9. Fusion peptides are visible for each of the subclones within the region delimited by the bracket and can be seen to be larger than the β-galactosidase control protein identified by the arrow in the control in Fig. 6A, lane c. Identical protein patterns were blotted onto nitrocellulose and probed with each of the selected antibodies (Figs. 6B, C and D). Antibodies selected by subclone 1 reacted only with fusion peptide

```
1     ..............................GAATTCCTTGCCGATTTTAAATCTAGGGTGTAGAGTTAGAGAAGGAGGTGTCTTTACATATGTGTGTACATTATATGAATGGAATCATTAG

92    GAATCACGTGAAAATCAAGTGGATCTATGATAAAAACAGCTTATTTAATTAAATCAAGTAATAATTCATAAATCTATAAATGTGCTTGAAAAAAATCCTATTTTCAGTTAAGAGATCTAC

212   GACATTATGTGTCAGAAATAAACCATAATTTTTTTCCATTCGCTAATCTTTTACGGCCTCGAACGTGGACTTGTGGATTCTAATATTACAAACTCTTTCACTGTTCATCTAGCTTTTGGA

332   GGGGCAGTCGACTCCGGATCTTTCGAATCTATAATTCCAATTTCCCCATACCTATCATTAAATTAATAGAATATATTAAATACATTTTCAGTTCTTGTAAAATATTGTAACATATCCACA

452   ATG AAG TTT CTT ATT TTG CTA TTT AAC ATT TTA TGT TTG TTC CCA GTT TTG GCG GCA GAC AAC CAC GGT GTT GGT CCT CAA GGG GCC TCC
1     Met Lys Phe Leu Ile Leu Leu Phe Asn Ile Leu Cys Leu Phe Pro Val Leu Ala Ala Asp Asn His Gly Val Gly Pro Gln Gly Ala Ser

542   GGT GTA GAT CCT ATA ACT TTT GAT ATT AAT TCA AAT CAA ACA GGC CCT GCA TTT TTA ACG GCC GTA GAG ATG GCC GGT GTA AAA TAT CTT
31    Gly Val Asp Pro Ile Thr Phe Asp Ile Asn Ser Asn Gln Thr Gly Pro Ala Phe Leu Thr Ala Val Glu Met Ala Gly Val Lys Tyr Leu

632   CAA GTA CAA CAT GGC TCT AAC GTA AAT ATC CAT AGA CTT GTT GAA GGG AAC GTT GTA ATT TGG GAG AAT GCA TCA ACA CCG TTG TAT ACC
61    Gln Val Gln His Gly Ser Asn Val Asn Ile His Arg Leu Val Glu Gly Asn Val Val Ile Trp Glu Asn Ala Ser Thr Pro Leu Tyr Thr

722   GGC GCA ATC GTT ACT AAT AAT GAC GGG CCT TAT ATG GCC TAT GTT GAG GTA CTG GGT GAT CCA AAT CTA CAG TTT TTT ATA AAA TCA GGT
91    Gly Ala Ile Val Thr Asn Asn Asp Gly Pro Tyr Met Ala Tyr Val Glu Val Leu Gly Asp Pro Asn Leu Gln Phe Phe Ile Lys Ser Gly

812   GAT GCT TGG GTG ACC TTA AGT GAG CAT GAA TAT TTA GCT AAA CTA CAG GAA ATC AGA CAA GCA GTT CAT ATT GAA TCA GTG TTT TCT CTT
121   Asp Ala Trp Val Thr Leu Ser Glu His Glu Tyr Leu Ala Lys Leu Gln Glu Ile Arg Gln Ala Val His Ile Glu Ser Val Phe Ser Leu

902   AAC ATG GCC TTC CAA CTG GAA AAT AAC AAA TAT GAA GTT GAA ACG CAC GCT AAA AAT GGA GCA AAT ATG GTG ACA TTT ATA CCA AGG AAT
151   Asn Met Ala Phe Gln Leu Glu Asn Asn Lys Tyr Glu Val Glu Thr His Ala Lys Asn Gly Ala Asn Met Val Thr Phe Ile Pro Arg Asn

992   GGA CAT ATT TGC AAA ATG GTA TAT CAT AAA AAT GTC AGA ATA TAC AAA GCA ACG GGC AAT GAC ACT GTT ACC TCT GTG GTG GGA TTT TTT
181   Gly His Ile Cys Lys Met Val Tyr His Lys Asn Val Arg Ile Tyr Lys Ala Thr Gly Asn Asp Thr Val Thr Ser Val Val Gly Phe Phe

1082  AGG GGT TTG AGG TTA TTG CTT ATT AAC GTT TTT AGT ATT GAC GAT AAT GGA ATG ATG AGT AAC AGA TAC TTT CAA CAT GTA GAT GAC AAA
211   Arg Gly Leu Arg Leu Leu Leu Ile Asn Val Phe Ser Ile Asp Asp Asn Gly Met Met Ser Asn Arg Tyr Phe Gln His Val Asp Asp Lys

1172  TAT GTT CCT ATT AGT CAA AAA AAC TAT GAG ACT GGT ATA GTT AAA TTA AAA GAC TAT AAG CAT GCT TAT CAT CCT GTA GAC TTG GAT ATT
241   Tyr Val Pro Ile Ser Gln Lys Asn Tyr Glu Thr Gly Ile Val Lys Leu Lys Asp Tyr Lys His Ala Tyr His Pro Val Asp Leu Asp Ile

1262  AAA GAT ATT GAT TAC ACC ATG TTC CAT TTG GCT GAT GCC ACT TAT CAT GAA CCT TGT TTC AAA ATT ATT CCA AAC ACA GGG TTT TGT ATA
271   Lys Asp Ile Asp Tyr Thr Met Phe His Leu Ala Asp Ala Thr Tyr His Glu Pro Cys Phe Lys Ile Ile Pro Asn Thr Gly Phe Cys Ile

1352  ACT AAG CTT TTT GAC GGT GAC CAA GTG CTC TAT GAG AGT TTT AAT CCC TTA ATT CAC TGT ATC AAT GAA GTA CAT ATT TAT GAT AGA AAT
301   Thr Lys Leu Phe Asp Gly Asp Gln Val Leu Tyr Glu Ser Phe Asn Pro Leu Ile His Cys Ile Asn Glu Val His Ile Tyr Asp Arg Asn

1442  AAC GGA TCA ATT ATT TGT CTA CAT CTA AAC TAC AGC CCA CCA TCC TAC AAA GCA TAC CTT GTT CTA AAA GAT ACT GGA TGG GAG GCT ACA
331   Asn Gly Ser Ile Ile Cys Leu His Leu Asn Tyr Ser Pro Pro Ser Tyr Lys Ala Tyr Leu Val Leu Lys Asp Thr Gly Trp Glu Ala Thr

1532  ACG CAC CCA TTA CTG GAA GAA AAG ATT GAA GAA TTG CAG GAC CAA AGA GCT TGC GAA TTG GAT GTA AAT TTT ATA AGT GAC AAA GAC CTT
361   Thr His Pro Leu Leu Glu Glu Lys Ile Glu Glu Leu Gln Asp Gln Arg Ala Cys Glu Leu Asp Val Asn Phe Ile Ser Asp Lys Asp Leu

1622  TAT GTG GCA GCA TTG ACA AAC GCG GAT CTT AAT TAC ACA ATG GTA ACC CCT AGA CCG CAT CGC GAT GTT ATT AGG GTA TCA GAT GGA AGT
391   Tyr Val Ala Ala Leu Thr Asn Ala Asp Leu Asn Tyr Thr Met Val Thr Pro Arg Pro His Arg Asp Val Ile Arg Val Ser Asp Gly Ser

1712  GAA GTG TTG TGG TAT TAT GAA GGT CTT GAT AAT TTT CTC GTT TGT GCA TGG ATT TAT GTT AGT GAC GGT GTT GCA TCA CTT GTT CAT TTA
421   Glu Val Leu Trp Tyr Tyr Glu Gly Leu Asp Asn Phe Leu Val Cys Ala Trp Ile Tyr Val Ser Asp Gly Val Ala Ser Leu Val His Leu

1802  AGG ATA AAA GAT CGT ATT CCA GCA AAT AAC GAT ATA TAC GTT CTC AAG GGT GAC CTA TAT TGG ACA AGA ATT ACA AAG ATA CAA TTT ACA
451   Arg Ile Lys Asp Arg Ile Pro Ala Asn Asn Asp Ile Tyr Val Leu Lys Gly Asp Leu Tyr Trp Thr Arg Ile Thr Lys Ile Gln Phe Thr

1892  CAA GAG ATA AAG AGG TTG GTC AAA AAA TCC AAG AAA AAA CTT GCT CCA ATT ACT GAA GAA GAC TCA GAT AAG CAC GAT GAA CCC CCA GAG
481   Gln Glu Ile Lys Arg Leu Val Lys Lys Ser Lys Lys Lys Leu Ala Pro Ile Thr Glu Glu Asp Ser Asp Lys His Asp Glu Pro Pro Glu

1982  GGA CCT GGA GCT TCA GGT TTG CCA CCC AAA GCG CCT GGA GAT AAA GAG GGA TCA GAA GGA CAT AAA GGA CCT AGT AAG GGA TCT GAT TCA
511   Gly Pro Gly Ala Ser Gly Leu Pro Pro Lys Ala Pro Gly Asp Lys Glu Gly Ser Glu Gly His Lys Gly Pro Ser Lys Gly Ser Asp Ser

2072  TCC AAG GAA GGT AAG AAA CCT GGA TCA GGT AAA AAA CCT GGA CCA GCA AGG GAA CAC AAA CCA AGC AAG ATA CCA ACA TTA TCC AAG AAG
541   Ser Lys Glu Gly Lys Lys Pro Gly Ser Gly Lys Lys Pro Gly Pro Ala Arg Glu His Lys Pro Ser Lys Ile Pro Thr Leu Ser Lys Lys

2162  CCT TCA GGA CCA AAA GAT CCA AAA CAT CCT AGG GAT CCT AAA GAA CCC AGG AAA TCT AAG TCG CCT AGG ACA GCA TCT CCT ACA AGG CGT
571   Pro Ser Gly Pro Lys Asp Pro Lys His Pro Arg Asp Pro Lys Glu Pro Arg Lys Ser Lys Ser Pro Arg Thr Ala Ser Pro Thr Arg Arg

2252  CCA TCA CCT AAA CTA CCC CAG TTA TCA AAG TTG CCT AAA TCA ACA TCT CCT AGA AGT CCT CCA CCA CCA ACA AGG CCT TCA TCA CCT GAA
601   Pro Ser Pro Lys Leu Pro Gln Leu Ser Lys Leu Pro Lys Ser Thr Ser Pro Arg Ser Pro Pro Pro Pro Thr Arg Pro Ser Ser Pro Glu

2342  CGT CCT GAG GGA ACT AAG ATA ATC AAA ACA TCT AAA CCA CCA TCT CCT AAA CCA CCG TTT GAT CCA TCA TTC AAG GAG AAA TTC TAT GAC
631   Arg Pro Glu Gly Thr Lys Ile Ile Lys Thr Ser Lys Pro Pro Ser Pro Lys Pro Pro Phe Asp Pro Ser Phe Lys Glu Lys Phe Tyr Asp

2432  GAT TAC AGT AAA GCT GCA AGT AGG TCT AAG GAA ACT AAG ACG ACA GTT GTA TTG GAT GAA TCA TTT GAA TCA ATA CTT AAG GAA ACT CTC
661   Asp Tyr Ser Lys Ala Ala Ser Arg Ser Lys Glu Thr Lys Thr Thr Val Val Leu Asp Glu Ser Phe Glu Ser Ile Leu Lys Glu Thr Leu
                                              [----- SUBCLONE 1 ----------------------------------------------------------

2522  CCT GAG ACA CCT GGA ACT CCC TTT ACC ACT CCT AGA CCT GTG CCA CCA AAA CGT CCC AGG ACA CCT GAA AGT CCC TTT GAA CCA CCT AAG
691   Pro Glu Thr Pro Gly Thr Pro Phe Thr Thr Pro Arg Pro Val Pro Pro Lys Arg Pro Arg Thr Pro Glu Ser Pro Phe Glu Pro Pro Lys
      ^^^
```

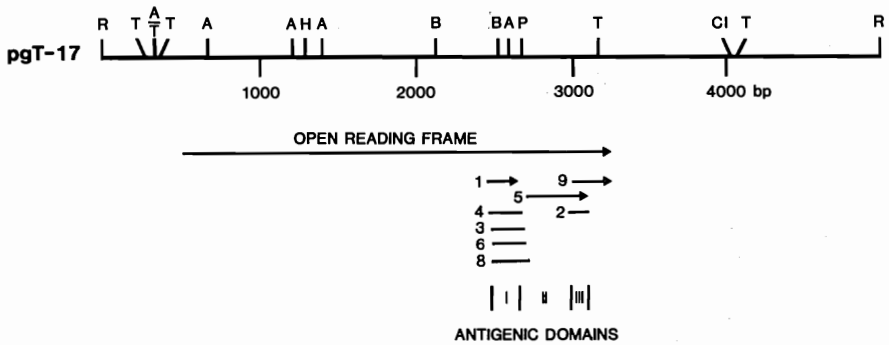

Fig. 4. Restriction map of the *T. parva parva* (Muguga) DNA insert in clone pgT-17. Restriction enzyme sites are designated as follows: *Eco*RI, R; *Bam*HI, B; *Hin*dIII, H; *Pst*I, P; *Cla*I, Cl; *Acc*I, A; *Taq*I, T. The open reading frame is delineated beneath the map. The eight subclones that define the antigenic regions of the open reading frame are aligned below the map and are designated by subclone numbers. The direction of transcription of subclones 1, 5 and 9 in λgt11 is indicated by the arrowheads. Designated by roman numerals at the bottom of the figure are the boundaries delimiting the three antigenic domains defined by the subclones.

encoded by this subclone. Antibody selected by peptides encoded by subclones 9 (Fig. 6C) and 5 (Fig. 6D) reacted with both fusion peptides 9 and 5, but not with fusion peptide expressed by subclone 1. The cross-reactivity between peptides 5 and 9 could result from common epitopes encoded within the overlapping sequences present in the two clones that does not include the proline-rich repeat. No cross-reactivity occurs between fusion products from subclones 1 and 9 (antigenic domains I and III), despite the presence of similar proline-rich repeat regions in both peptides.

The 90- and 85-kDa proteins lack determinants found in the 104-kDa antigen. The antigens recognised by whole C16 antiserum on Western blots of sporozoite extracts prepared from uninfected (U) and DE52-purified *Theileria*-infected (I) tick salivary gland extracts are shown in Fig. 7A. Domain-specific antibodies reactive with each of the three fusion peptides were next employed to probe identical Western blots of uninfected and infected salivary extracts. As shown in Fig. 7, antibodies selected by subclone 1 (Fig. 7B), subclone 5 (Fig. 7C) and subclone 9 (Fig. 7D) all reacted with the 104-kDa antigen. This confirms that each of the antigenic domains identified is expressed within the 104-kDa antigen. Each set of selected antibodies also reacted with the 35-kDa antigen, showing that this peptide also expresses determinants found in all three of the antigenic domains. With respect to the antigens of approximately 90 and 85 kDa, antibodies selected by subclone 9 did not react with either the 90- or the 85-kDa antigen, while antibodies selected by subclone 5 recognised the 90-kDa but not the 85-kDa antigen. Only antibodies selected by subclone 1 reacted with both of the antigens. Western blots were also carried out using antisera raised in calves against purified fusion peptides encoded by subclones 1 and 9 (Fig. 7, panels E and F, respectively) which confirmed the results obtained

SUBCLONE 1	SUBCLONE 9
Thr Leu Pro	**Arg Arg** Pro Pro
Glu **Thr** Pro	**Lys Lys** Pro
Gly **Thr** Pro	Ser **Lys** Ser
Phe **Thr Thr**	Pro **Arg** Pro
Pro Arg Pro	Ser **Lys** Pro
Val Pro Pro	**Lys Lys** Pro
Lys Arg Pro	**Lys Lys** Pro
Arg **Thr** Pro	
Glu **Ser** Pro	
Phe Glu Pro Pro	
Lys Asp Pro	
Asp **Ser** Pro **Ser**	
Thr Ser Pro	

Fig. 5. Vertical alignment of proline residues from the repeated amino acid sequences of subclones 1 and 9. Sequence encoded by subclone 5 also contains a high proportion of hydroxylated amino acids, serine and threonine (bold-face type), while highly charged amino acids, arginine and lysine (bold-face type) predominate in the amino acid sequence encoded by subclone 9.

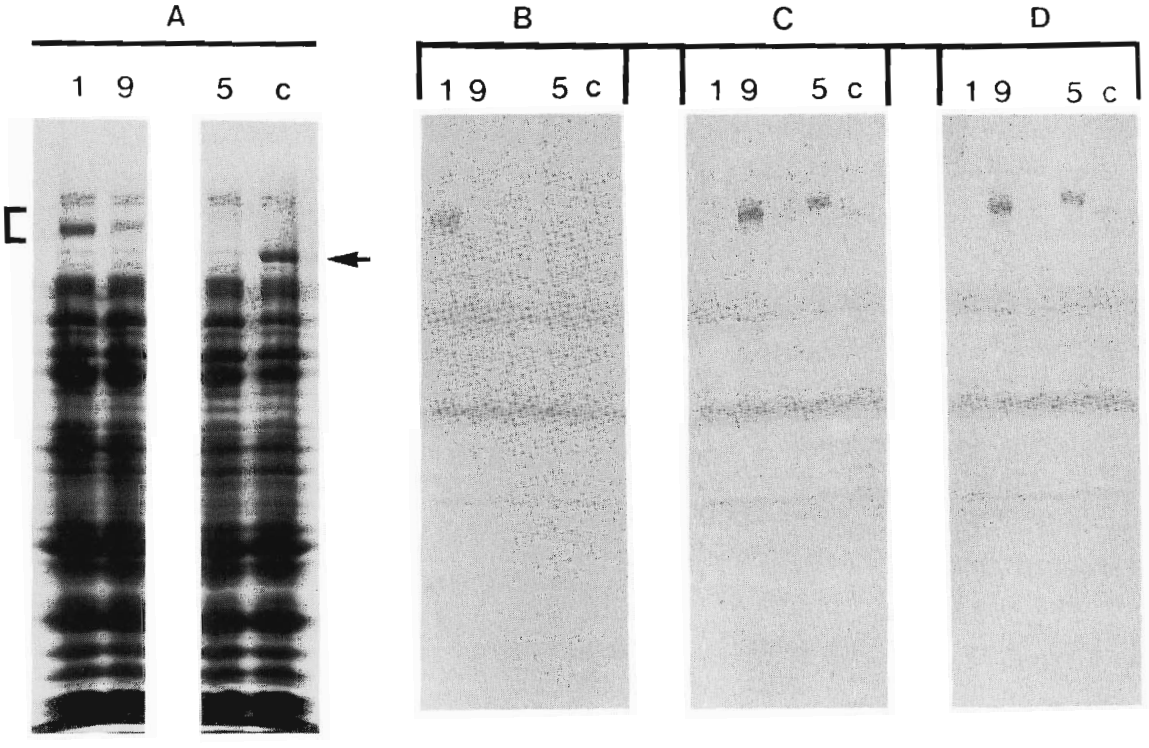

Fig. 6. Panel A shows Coomassie Blue 250-stained protein patterns made from lysates of induced lysogens of subclones 1, 9 or 5 resolved on a 7.5% polyacrylamide gel. Fusion peptides are located within the bracketed region of the gel pattern. Control lane c contains a lysate made from lysogenised λgt11 vector. The location of β-galactosidase is delineated by the arrow. Identical patterns were blotted onto nitrocellulose and probed with antibodies selected by peptides encoded by subclones 1 (panel B), 9 (panel C) and 5 (panel D).

with the antibodies selected by each of these two clones. The fusion peptide encoded by subclone 5 did not elicit an antibody response in the one calf that was immunised.

The 104-kDa antigen occurs in the microneme/rhoptry complexes. Antisera raised against subclone 1 were used to localise the antigen in sections of *T. parva* sporozoites. Fig. 8 shows electron micrographs of embedded and sectioned sporozoites probed with anti-subclone 1 serum. Antibody reactivity was revealed using protein A-gold conjugate. The signals appear to be localised in only a portion of the organelles comprising the microneme/rhoptry complex. Sporozoites incubated with unrelated serum and protein A-gold alone, had no gold particles associated with them (data not shown).

The 104-kDa antigen is recognised by recovery serum. Sporozoite extracts were also probed with recovery serum taken about three months after challenge of a cow (B641) that had previously been immunized against theileriosis by the infection and treatment procedure. Shown in Fig. 9 is a comparison of the antigen profiles detected by anti-sporozoite serum, C16, and by immune serum, B641. These results show that the 104-kDa antigen and antigenically related proteins of 90 and 85 kDa are recognised by this serum, along with an unrelated molecule of approximately 82 kDa. Similar results have been obtained with sera from immune cattle that were subject to natural field challenge (data not shown), indicating that immunological reactivity to epitopes expressed on the 104-kDa antigen is maintained in animals immunised by the infection and treatment method.

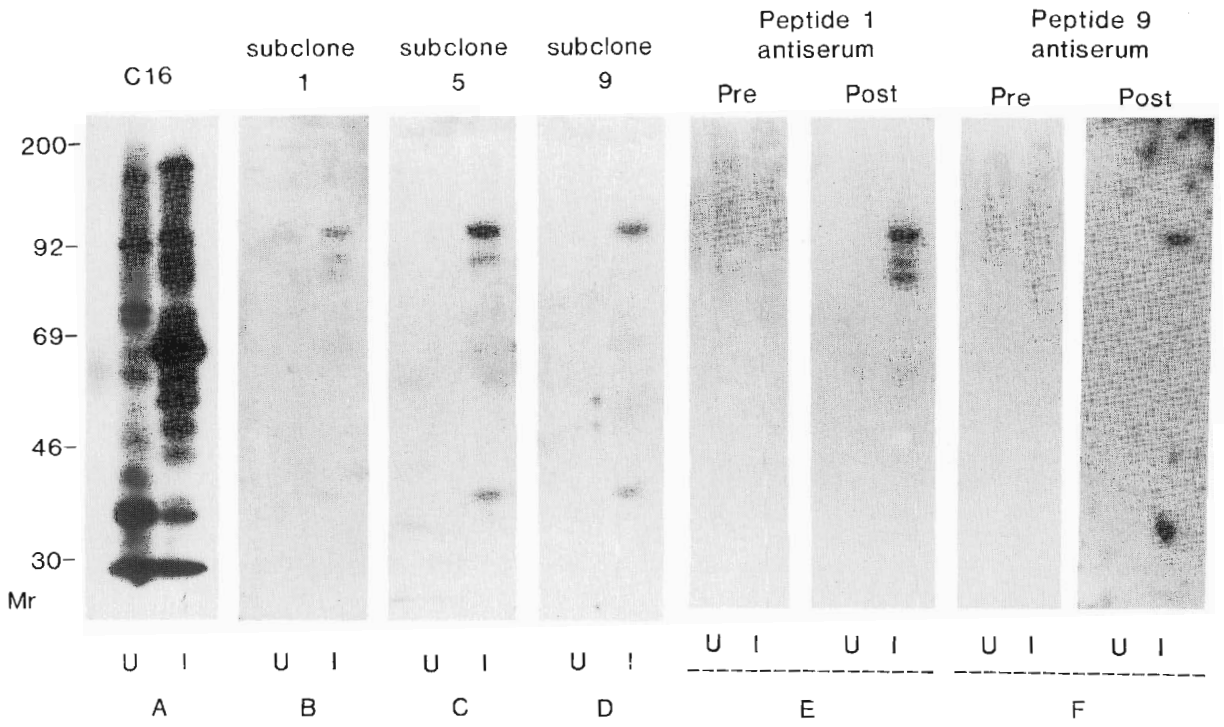

Fig. 7. Western blots of sporozoite extracts made from uninfected (U) and *Theileria*-infected (I) tick salivary glands probed with antibodies reactive against the fusion products of subclones 1, 5 and 9. Panel A shows patterns produced by probing Western blots with whole C16 antiserum. Identical blots were probed with antibodies obtained either by affinity-purification from C16 antiserum using fusion products of subclones 1 (7B), 5 (7C) and 9 (7D) or by immunisation of cattle with purified fusion products from subclones 1 (7E) and 9 (7F). In cattle immunized with fusion peptide from subclone 1 and 9, pre-immune serum was tested for lack of reactivity to theilerial antigens (7E and F, respectively). Antigen-antibody complexes were made visible using ^{125}I-labelled protein G. Molecular mass markers are shown on the left.

Discussion

A λgt11 recombinant phage clone has been identified that encodes determinants reactive with antisera taken from calves immunised with freeze-thawed extracts of semi-purified sporozoites. These determinants reside on proteins of approximately 104, 90, 85 and 35 kDa. On further characterisation of the cloned DNA fragment, we found it to encode sequences that are expressed during sporozoite maturation. On sequencing a portion of the DNA insert, a 2772-bp open reading frame was identified that was not fused to the vector's *lacZ* gene. The conceptual translation product would include 924 amino acid residues that would have an approximate size of 104 kDa. Thus, there is no evidence for the presence of introns in this gene. Hydrophobic stretches occur at the beginning and end of the open reading frame, suggesting a possible membrane affiliation for the protein.

The sequence encoded by clone λgTpS-17 included 445 bp of presumptive promoter region, extending upstream to the *Eco*RI linker. The occurrence of this sequence would have prevented production of a fusion product by introducing a large number of termination codons between the β-galactosidase initiation codon and that of the 104-kDa antigen. This would explain why we were not able to produce fusion protein products from lysogens of λgTpS-17. It would also indicate that some of the transcriptional or translational control of this cloned sequence was mediated by the theilerial sequence, which was apparently able to function within the bacterial host. It was also found that expression of the 104-kDa antigen, al-

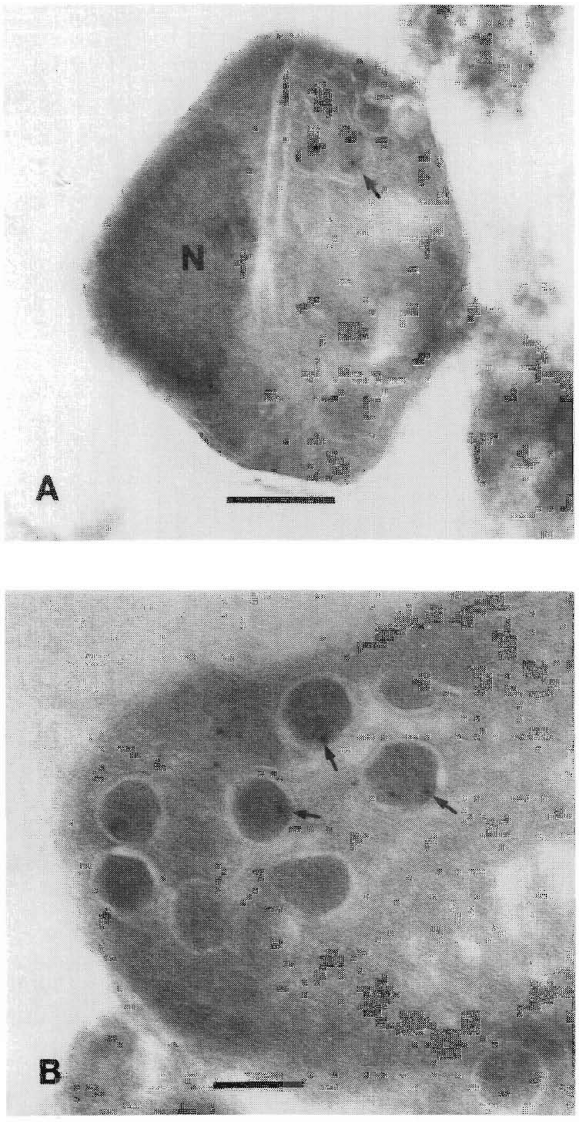

Fig. 8. Electron micrographs showing the immunolocalisation of the 104-kDa antigen in sporozoites. The antibodies, made visible by protein A-gold, specifically label the microneme/rhoptry complexes. In panel A, the labelled complexes are arrowed. Protein A-gold was not seen over the rest of the sporozoite, including the nucleus (N). Bar, 0.2 μM. Panel B is a higher magnification of the microneme/rhoptry complexes of another labelled sporozoite. Some organelles are labelled with protein A-gold (arrows) while others are not. Bar, 0.1 μM.

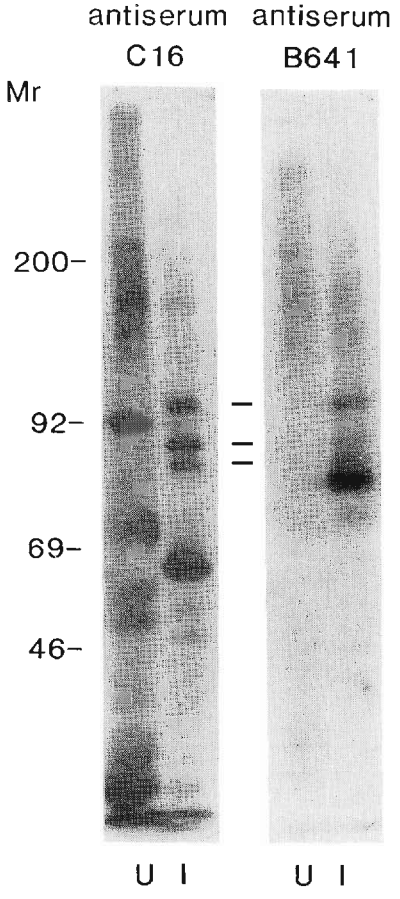

Fig. 9. Western blot of sporozoite extracts made from uninfected (U) and *Theileria*-infected (I) tick salivary glands probed with C16 antiserum and with serum, B641, taken from a calf immunised against *T. parva parva* (Muguga) by the infection and treatment method. Molecular mass markers are indicated on the left.

though reduced, was detectable on immunoscreens without treatment of filters with IPTG. This is further evidence that the gene was not entirely under the control of the β-galactosidase gene promoter. Additionally, it was not possible to detect the intact 104-kDa protein on Western blots of denaturing SDS-PAGE gels, and we assume that this indicates that rapid and extensive degradation of the molecule was occurring during growth and synthesis in culture or during handling and preparation of the bacterial lysates.

The cloned sequence coding for this 104-kDa antigen was subfragmented and re-expressed in λgt11 in an attempt to localise those regions that are antigenic. Three distinct antigenic domains were defined which include an unknown number of individual antigenic determinants. The pre-

dicted amino acid sequence of two domains shows structural similarity to antigenic regions found in other protozoan parasite antigens. Domains I and III are both characterised by an unusually large number of proline residues which are repeated at approximately three amino acid intervals. Repetitive sequences have been reported in *Plasmodium* and trypanosome parasite antigens, although they have not previously been seen in rhoptry proteins [25–32]. It has previously been shown that repetitive elements are often immunodominant. It has been suggested that apparent immunodominance results from repetitiveness of the epitope and that such sequences are often preferentially identified as a result of their resistance to denaturation during immunoscreening procedures [34–37]. In the case of the 104-kDa antigen, subclone 5 expresses determinants not related to the proline-rich repeats of clones 1 and 9. This is apparent from the reactivity of antibodies selected by clone 5 with the 90-kDa antigen. This reactivity does not occur with antibodies selected by clone 9, and is therefore not encoded in the regions that are shared by the two clones. It therefore follows that immunodominance is not as pronounced with these repetitive elements as it is with antigens described for other parasites.

Antibodies selected by the fusion peptides from subclones 1, 5 or 9 were employed to test for cross-reactivity of the epitopes found in the three antigenic domains. Antibodies selected by fusion peptides encoded in subclones 5 and 9 (antigenic domains II and III), cross-reacted with each other, possibly as a result of the common sequence shared between them. Neither set of selected antibodies reacted with fusion peptide from subclone 1 (antigenic domain I). The lack of cross-reactivity between peptides encoded by subclones 9 and 1 suggests that the proline-rich repeated regions in the 104-kDa antigen also do not

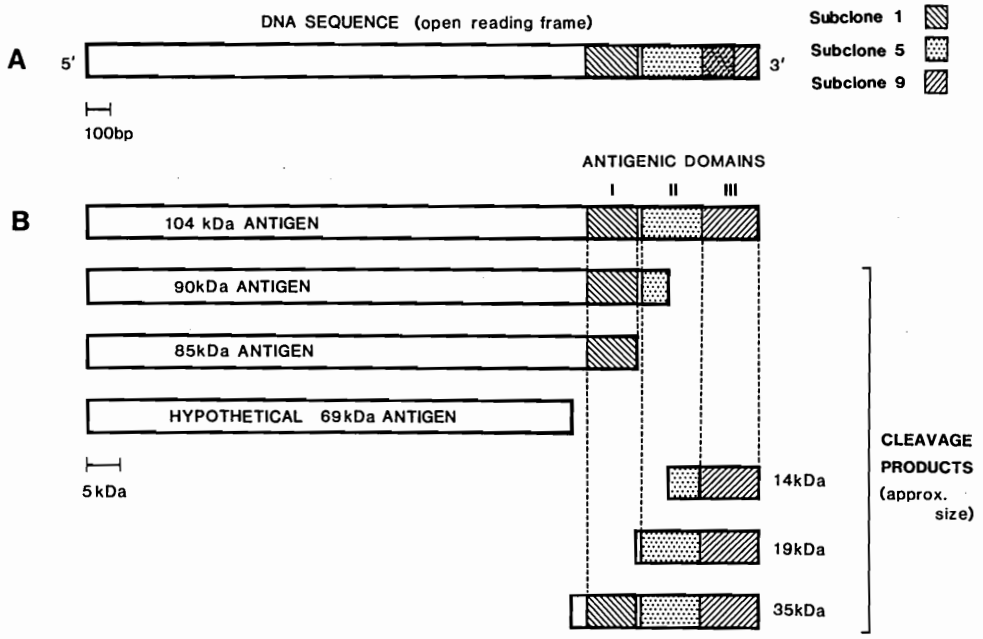

Fig. 10. Diagrammatic representation of the organisation of the gene sequence encoding the 104-kDa sporozoite antigen and the proposed derivation of the 90-, 85- and 35-kDa antigens from the 104-kDa antigen. In diagram A are shown the locations of the three subclones (1, 5 and 9) that encode sequences at the 3′ end of the open reading frame of clone pgT-17. These sequences encode the determinants expressed in antigenic domains I, II and III. In diagram B are shown the proposed relationship among the four antigens. Cleavage of a 14-kDa fragment from the 104-kDa antigen produces a 90-kDa antigen lacking all of antigenic domain III and portions of antigenic domain II. A 19-kDa cleavage product would produce an 85-kDa antigen and remove all determinants expressed in domains II and III, leaving only those of domain I. The 35-kDa antigen could be produced by cleavage from the -COOH terminus of the 104-kDa antigen, and would include all of the determinants expressed within the three antigenic domains. The occurrence of this antigen would predict the existence of an undetected protein of about 69-kDa.

exhibit the cross-reactive characteristics reported for other parasite antigens containing repetitive elements. The lack of cross-reactivity between the two fusion peptides may result from the different composition of the non-proline residues found in the two repetitive tracts, one containing many basic and the other many hydroxylated residues. It might also be attributable to one or both of the repeats being non-antigenic, as the data in no way prove the antigenicity of these two repetitive amino acid sequences.

Western blots of extracts of sporozoites were probed with antibodies selected by each of the three subclones studied. The results provide insight into the nature of the relationship between the 104-kDa protein and the smaller immunologically related proteins of 90, 85 and 35 kDa. The fact that antibody selected by the fusion product of subclone 9, which forms the presumptive carboxyl end of the 104-kDa antigen, does not cross-react with the 90- or 85-kDa proteins, suggests cleavage of this end of the molecule to produce the smaller antigens. Antibody selected by subclone 5 fusion peptide reacted with the 104-kDa, 90-kDa and 35-kDa antigens. The reactivity of these antibodies with the 90-kDa antigen suggests that the initial cleavage reaction separated the protein sequence common to subclones 5 and 9 from those portions exclusively encoded by subclone 5. The fact that amino acid sequences exclusive to subclone 5 fusion peptide are recognised in the smaller antigen also proves that additional epitopes are encoded within the DNA of subclone 5 which are not shared with subclone 9. Antibody to subclone 1 fusion peptide reacted with the 104-, 90-, 85- and 35-kDa antigens, thus indicating that this subclone encodes antigenic regions that are sufficiently far upstream of the carboxyl terminus of the protein to remain uncleaved from the amino terminal sequence through the initial two cleavages. The reactivity of the 35-kDa antigen with all of the selected antibodies can be explained in one of two ways. Cleavage of 35 kDa of the COOH terminus of the molecule could be occurring. This would encompass the entire antigenic region of the protein, leaving a 69-kDa protein. Although C16 antiserum does react with a 69-kDa antigen, antibodies eluted from the 69-kDa region of Western blots have shown no cross-reactivity with the 104-, 90- or 85-kDa antigens (K.P.I., unpublished observation). This would indicate that, if a 69-kDa molecule exists, it does not react with C16 antiserum, and is therefore non-antigenic. It could also be postulated that a distinct antigen of 35 kDa exists that contains epitopes that are cross-reactive with all three sets of selected antibodies. This would obviate the need to propose the existence of a non-antigenic 69-kDa molecule. The proposed relationship amongst the four antigens also predicts the occurrence of cleavage products of approximately 14 and 19 kDa, which would not be resolved on the polyacrylamide gels employed for these experiments. However, previous results obtained by one of us (A.J.M., unpublished observation) have identified low-molecular-weight species that are reactive with C16 antiserum. The proposed derivation of the three proteins by cleavage of the antigenic C-terminal end of the 104-kDa antigen is illustrated diagrammatically in Fig. 10. An alternative explanation to this proposed relationship would be the smaller antigens derive from shortened mRNA molecules, perhaps resulting from premature termination of transcription. However, this would have resulted in the appearance of several species of mRNA on Northern blots of *Theileria*-infected tick salivary glands, which were not seen. Additionally, each antigen could derive from a unique gene sequence. However, as only one signal was identified on a Southern blot probed with clone pgT-17, such genes would have to vary sufficiently in sequence to be non-homologous with pgT-17 while still expressing common antigenic determinants. Although the latter can not be entirely ruled out, we feel it to be unlikely.

The presence of the 104-kDa antigen within only a portion of the vesicles of the microneme/rhoptry complex may indicate reduced antibody accessibility of some sites, or biochemical diversity within the complex. Further studies, aimed at identifying the antigen in other life-cycle stages and in characterising its functional role within the rhoptries, are now under way.

Acknowledgements

We would like to thank Dr. Ivan Morrison for providing recovery antiserum, B641, and Ms. Elke Gobright and Mr. Kibe Kanyi M. for technical assistance.

References

1 Webster, P., Dobbelaere, D.A.E. and Fawcett, D.W. (1985) Entry of sporozoites of *Theileria parva* into bovine lymphocytes: immunoelectron microscopic observations. Eur. J. Cell Biol. 36, 157–162.

2 Fawcett, D.W., Doxsey, S., Stagg, D.A. and Young, A.S. (1982) The entry of sporozoites of *Theileria parva* into bovine lymphocytes in vitro. Eur. J. Cell Biol. 27, 10–21.

3 Kilejian, A.A. (1974) A unique histidine-rich polypeptide from the malaria parasite, *Plasmodium lophurae*. J. Biol. Chem. 249, 4650–4655.

4 Roger, N., Dubremetz, J.-F., Delplace, P., Fortier, B., Tronchin, G. and Vernes, A. (1988) Characterisation of a 225-kDa rhoptry protein of *Plasmodium falciparum*. Mol. Biochem. Parasitol. 27, 135–142.

5 Uni, S., Masuda, A., Stewart, J., Igarashi, I., Nussenzweig, R. and Aikawa, M. (1987) Ultrastructural localisation of the 150/130 kDa antigens in sexual and asexual blood stages of *Plasmodium falciparum*-infected human erythrocytes. Am. J. Trop. Med. Hyg. 36, 481–488.

6 Cooper, J.A., Ingram, L.T., Bushell, G.R., Fardoulys, C.A., Stenzel, D., Schofield, L. and Saul, A.J. (1988) The 140/130/105 kilodalton protein complex in the rhoptries of *Plasmodium falciparum* consists of discrete polypeptides. Mol. Biochem. Parasitol. 29, 251–260.

7 Campbell, G.H., Miller, L.H., Hudson, D., Franco, E.L. and Andrysiak, P.M. (1984) Monoclonal antibody characterisation of *Plasmodium falciparum* antigens. Am. J. Trop. Med. Hyg. 33, 1051–1054.

8 Holder, A.A., Freeman, R.R., Uni, S. and Aikawa, M. (1985) Isolation of a *Plasmodium falciparum* rhoptry protein. Mol. Biochem. Parasitol. 14, 293–303.

9 Clark, J.T., Anand, R., Akoglu, T. and McBride, J.S. (1987) Identification and characterisation of proteins associated with the rhoptry organelles of *Plasmodium falciparum* merozoites. Parasitol. Res. 73, 425–434.

10 Bushell, G.R., Ingram, L.T., Fardoulys, C.A. and Cooper, J.A. (1988) An antigenic complex in the rhoptries of *Plasmodium falciparum*. Mol. Biochem. Parasitol. 28, 105–112.

11 Holder, A.A. and Freeman, R.R. (1981) Immunization against blood-stage rodent malaria using purified parasite antigens. Nature 294, 361–364.

12 Oka, M., Aikawa, M., Freeman, R.R., Holder, A.A. and Fine, E. (1984) Ultrastructural localisation of protective antigens of *Plasmodium yoelii* merozoites by the use of monoclonal antibodies and ultrathin cryomicrotomy. Am. J. Trop. Med. Hyg. 33, 342–346.

13 Sadak, A., Taghy, Z., Fortier, B. and Dubremetz, J. (1988) Characterisation of a family of rhoptry proteins of *Toxoplasma gondii*. Mol. Biochem. Parasitol. 29, 203–211.

14 Schofield, L., Bushel, G.R., Cooper, J.A., Saul, A.J., Upcroft, J.A. and Kidson, C. (1986) A rhoptry antigen of *Plasmodium falciparum* contains conserved and variable epitopes recognized by inhibitory monoclonal antibodies. Mol. Biochem. Parasitol. 18, 183–195.

15 Schwartzman, J.D. (1986) Inhibition of a penetration-enhancing factor of *Toxoplasma gondii* by monoclonal antibodies specific for rhoptries. Infect. Immun. 21, 760–764.

16 Dolan, T.T., Radley, D.E., Brown, C.G.D., Cunningham, M.P., Morzaria, S.P. and Young, A.S. (1980) East Coast fever: further studies on the protection of cattle immunized with a combination of theileria strains. Vet. Parasitol. 6, 325–332.

17 Taylor, D.W., Cordingley, J.C. and Butterworth, A.E. (1984) Immunoprecipitation of surface antigen precursors from *Schistosoma mansoni* messenger RNA in vitro translation products. Mol. Biochem. Parasitol. 10, 305–318.

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

19 Meinkoth, J. and Wahl, G. (1984) Hybridization of nucleic acids immobilized on solid supports. Anal. Biochem. 138, 267–284.

20 Young, R.A., Mehra, V., Sweetser, D., Buchanan, T., Clark-Curtiss, J., Davis, R.W. and Bloom, B.R. (1985) Genes for the major protein antigens of the leprosy parasite *Mycobacterium leprae*. Nature 316, 450–452.

21 Ozaki, L.S., Mattei, D., Jendoubi, M., Druihle, P., Blisnick, T., Guillote, M., Puijalon, O. and Pereira da Silva, L. (1986) Plaque antibody selection: rapid immunological analysis of a large number of recombinant phage clones positive to sera raised against *Plasmodium falciparum* antigens. J. Immunol. Methods 89, 213–219.

22 Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680–685.

23 Ullmann, A. (1984) One-step purification of hybrid proteins which have β-galactosidase activity. Gene 29, 27–31.

24 Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5467.

25 Coppel, R.L., Bianco, A.E., Culvenor, J.G., Crewther, P.E., Brown, G.V., Anders, R.F. and Kemp, D.J. (1987) A cDNA clone expressing a rhoptry protein of *Plasmodium falciparum*. Mol. Biochem. Parasitol. 25, 73–81.

26 Dame, J.B., Williams, J.L., McCutchan, T.F., Weber, J.L., Wirtz, R.A., Hockmeyer, W.T., Maloy, W.L., Haynes, J.D., Schneider, I., Roberts, D., Sanders, G.S., Reddy, E.P., Diggs, C.L. and Miller, L.H. (1984) Structure of the gene encoding the immunodominant antigen on the sporozoite of the human malaria parasite *Plasmodium falciparum*. Science 225, 593–599.

27 Coppel, R.L., Cowman, A.F., Lingelbach, K.R., Brown,

G.V., Saint, R.B., Kemp, D.J. and Anders, R.F. (1983) Isolate specific S-antigen of *Plasmodium falciparum* contains a repeated sequence of eleven amino acids. Nature 306, 751–756.

28 Roditi, I., Carrington, M. and Turner, M. (1987) Expression of a polypeptide containing a dipeptide repeat is confined to the insect stage of *Trypanosoma brucei*. Nature 325, 272–274.

29 Cowman, A.F., Saint, R.B., Coppel, R.L., Brown, G.V., Anders, R.F. and Kemp, D.J. (1985) Conserved sequences flank variable tandem repeats in two S-antigen genes of *Plasmodium falciparum*. Cell 40, 775–783.

30 Ozaki, L.S., Svec, P., Nussenzweig, R.S., Nussenzweig, V. and Godson, G.N. (1983) Structure of the *Plasmodium knowlesi* gene coding for the circumsporozoite protein. Cell 34, 815–822.

31 Enea, V., Ellis, J., Zavala, F., Arnot, D.E., Asavanich, A., Musada, A., Quakyi, I. and Nussenzweig, R.S. (1984) DNA cloning of *Plasmodium falciparum* circumsporozoite gene: amino acid sequence of repetitive epitope. Science 225, 628–630.

32 McCutchan, T.F., Lal, A.A., De la Cruz, V., Miller, L.H., Maloy, W.L., Charoenvit, Y., Beaudoin, R.L., Guerry, P., Wistar, R., Hoffman, S.L., Hockmeyer, W.T., Collins, W.E. and Wirth, D. (1985) Sequence of the immunodominant epitope for the surface protein on sporozoites of *Plasmodium vivax*. Science 230, 1381–1383.

33 Sharma, S., Svec, P., Mitchell, G.H. and Godson, G.N. (1985) Diversity of the circumsporozoite genes from two strains of the malarial parasite *Plasmodium knowlesi*. Science 229, 779–782.

34 Coppel, R.L., Cowman, A.F., Anders, R.F., Bianco, A.E., Saint, R.B., Lingelbach, K.R., Kemp, D.J. and Brown, G.V. (1984) Immune sera recognize on erythrocytes a *Plasmodium falciparum* antigen composed of repeated amino acids. Nature 310, 789–793.

35 Anders, R.F., Shi, P.T., Brown, G.V., Stahl, H.D., Favaloro, J., Bianco, A.E., Crewther, P.E., Coppel, R.L., Leach, S.J. and Kemp, D.J. (1985) Immune recognition of repeat structures in malaria antigens. In: Immune Recognition of Protein Antigens (Laver, W.G. and Air, G.M., eds.), pp. 193–197. Current Communications in Molecular Biology, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

36 Kemp, D.J., Coppel, R.L., Cowman, A.F., Saint, R.B., Brown, G.V. and Anders, R.F. (1983) Expression of *Plasmodium falciparum* blood-stage antigens in *Escherichia coli*: detection with antibodies from immune humans. Proc. Natl. Acad. Sci. USA 80, 3787–3791.

37 Zavala, F., Tam, J.P., Hollingdale, M.R., Cochrane, A.H., Quakyi, I., Nussenzweig, R.S. and Nussenzweig, V. (1985) Rationale for development of a synthetic vaccine against *Plasmodium falciparum* malaria. Science 228, 1436–1440.