DNA encoding an antigen of 101,000 apparent molecular weight from the human malaria parasite *Plasmodium falciparum* was cloned and sequenced. Genomic DNA from the Camp strain covering the complete coding region along with cDNA from the FCR3 strain covering 81% of the coding region were obtained. The cloned DNA specified a full-length protein of 743 amino acids which included two tandemly repeated regions, one near the amino terminus containing eight hexapeptide repeats of sequence TND, and the second near the carboxyl terminus containing primarily KE and KEE repeats. The latter repeated region is encoded by a 174-base stretch of mRNA containing only a single pyrimidine. Except for a putative leader sequence located at the amino terminus of the protein, the protein is hydrophilic and highly charged with a calculated isoelectric point of 5.6. Sequences from the Camp and FCR3 strains are very close and are also nearly identical to the partial cDNA sequence of the acidic basic repeated antigen (ABRA) protein from the FC27 strain (Stahl, H. D., Bianco, A. E., Crewther, R. F., Anders, R. F., Kyne, A. P., Coppel, R. L., Mitchell, G. F., Kemp, D. J., and Brown, G. V. (1986) *Mol. Biol. Med.* 3, 351–368). ABRA was previously shown to be located at the merozoite surface and in the parasitophorous vacuole. Because of its location and because it becomes complexed to merozoites when schizonts rupture in the presence of immune serum, ABRA is a candidate component of a malaria vaccine.

Malaria is still an enormous health problem in the tropical regions of the world. A malaria vaccine is urgently needed to augment the effectiveness of antimalarial drugs and insecticides. Because human malaria parasites cannot be obtained in sufficient quantity and purity to prepare vaccines directly, malaria vaccine development has relied on cloning and sequencing genes encoding parasite antigens (1). Unlike malaria sporozoites, which are coated with a single immunodominant protein, malaria blood stage merozoites express many antigens which are recognized by serum from immune individuals. Among the best blood stage vaccine candidates are antigens which appear to be located at the merozoite surface and are enriched in immune clusters of merozoites that form when schizonts rupture in the presence of immune serum (2-4). In this report, we describe the primary structure of one of these antigens, previously called p101 because it migrates on SDS-polyacrylamide gels with an apparent molecular weight of 101,000 (4).

Comparison of the p101 sequence to a partial cDNA sequence reported by Stahl and co-workers (5) revealed that p101 was identical to the acidic basic repeat antigen (ABRA). To avoid confusion, we have used ABRA to describe the p101 antigen throughout this paper.

**EXPERIMENTAL PROCEDURES**

**Culture of Parasites**—Cloned Camp strain *Plasmodium falciparum* parasites (Malaysia) were cultured and synchronized by double sorbitol treatment as described (6). Parasites were harvested by banding the schizont-infected erythrocytes on Percoll gradients (7).

**Gene Cloning and Sequencing**—Construction of the genomic DNA expression library using mung bean nuclease and screening of the library with monkey immune serum has been described (8, 9). The EcoRI/TaqI fragment from the 5' end of clone a65 (Fig. 1) was eluted from a preparative agarose gel using the GENECLEAN system (Bio 101, LaJolla, CA) and subcloned between the EcoRI and CiaI sites of pBR322. After amplification of the plasmid DNA and resolution of the insert, the fragment was nick-translated and hybridized (10) to Southern blots containing Camp strain genomic DNA digested with various restriction enzymes. In this way, genomic restriction enzyme sites were mapped within and about the ABRA gene. The same parasite DNA fragment was then used to screen a FCR3 strain (West Africa) cDNA library (11) and also to screen a minilibrary comprised of Camp strain genomic DNA EcoRI/BglII fragments, 3.5-3.9 kilobases in length ligated into pUC13 (see Fig. 1). A portion of the resulting ABRA cDNA clone from the 5' end of the clone to the EcoRI site was then isolated, labeled, and used to screen a minilibrary containing 2.8-3.2 kilobase TaqI/EcoRI genomic fragments ligated into pUC13.

Complete sequences of both strands of the Camp genomic DNA were determined by the Sanger dideoxy chain termination method using the Klenow fragment of DNA polymerase I and [α-32P]dATP (8). Subclones for sequencing were produced using T4 DNA polymerase as an exonuclease (12) or by subcloning restriction enzyme fragments into m13. Gaps were filled by using 17-18 base synthetic oligodeoxynucleotides as primers. The region around the polypurine/polypyrimidine segment (bp 2020-2193) required that reactions be carried out with deaza-dGTP at 50 °C (13). About 95% of both strands of the FCR3 cDNA clone were similarly sequenced; in the two regions where 1 The abbreviations used are: SDS, sodium dodecyl sulfate; ABRA, acidic basic repeat antigen; bp, base pairs; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid.

James L. Weber‡‡, Jeffrey A. Lyon‡, Ruth H. Wolff, Ted Hall‡, George H. Lowell‡, and Jeffrey D. Chulay‡

From the ‡Department of Immunology, Walter Reed Army Institute of Research, Washington, D. C. 20307-5100, the ¤Marshfield Medical Research Foundation, Marshfield, Wisconsin 54449, and the ¶Department of Bacterial Diseases, Walter Reed Army Institute of Research, Washington, D. C. 20307-5100

(Received for publication, February 9, 1988)

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) J03902.
only one strand was sequenced, the FCR3 sequence was identical to the Camp sequence. Sequences were analyzed with software from IBI (New Haven, CT) and BIONET (Intelligenetics, Mountain View, CA).

**Immunoblotting Experiments with Parasite Antigens**—Schizont antigens were fractionated by electrophoresis on SDS-polyacrylamide gels and electrobotted onto nitrocellulose filters (4). Filter strips were reacted with various antibody preparations, and after washing, the bound antibodies were detected using 125I-labeled secondary antibodies (9).

ABRA was purified from [14C]isoleucine-labeled *P. falciparum* schizonts cultured with protease inhibitors and extracted with 1% Triton X-100 (14). Antigens from 10^6 parasites were loaded onto a 3.5-ml column of monoclonal antibody 3DS (4) coupled to Affi-Gel 10 and washed with phosphate-buffered saline, pH 7.4. Bound ABRA was eluted with 0.1 M sodium citrate, pH 2.8, and neutralized with HEPES (85 mM final concentration) and NaOH. All buffers contained chymostatin (100 μg/ml) to inhibit proteolysis. As assessed by laser densitometry of autoradiographs after SDS-polyacrylamide gel electrophoresis, ABRA accounted for 4% of the total metabolically labeled antigens applied to the affinity column and was the major metabolically labeled antigen eluted from the column. A 2-ml aliquot containing 25% of the eluted radioactive antigen, or a control containing 2 ml of bovine serum albumin (10 μg/ml), was incubated overnight with a 1-inch diameter 0.2-μm nitrocellulose filter; more than 90% of radioactivity bound to the filter. The filters were then blocked with phosphate-buffered saline containing 3% bovine serum albumin and 0.3% Tween 20 and used to affinity purify antibodies (9).

Antibodies recognizing expression fusion proteins from αgt11 lysogens of both clone a95 and a circumsporozoite protein clone (15) were affinity-purified from monkey immune serum as described (9). A synthetic peptide with sequence NVVPTQSKKKNNET-VSGMDE based on the deduced amino acid sequence of ABRA (amino acids 516-537) was synthesized by Peninsula Laboratories, Inc. (Belmont, CA). The peptide was coupled to keyhole limpet hemocyanin using 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (16). Mice were immunized subcutaneously with the peptide carrier emulsified in Freund's complete adjuvant and boosted intravenously with the peptide carrier in saline.

**Immunoblotting Experiments with Fusion Proteins**—*Escherichia coli* strain Y1089 was lysogenized (17) with αgt11 bacteriophage without an insert or with the ABRA gene insert (clone a95). Lysogens were grown at 32°C in nutrient broth supplemented with 50 μg/ml ampicillin until the absorbance at 550 nm reached 0.5. The cultures were then incubated in a 44°C water bath for 20 min before adding 5% volumes of 100 mM isopropyl-β-D-thiogalactopyranoside or water and shaking for an additional 60 min at 37°C. *E. coli* cells were pelleted by centrifugation at 12,000 × g for 15 min. Pellets were resuspended in 2% culture volume of 50 mM Tris (pH 8.0), 150 mM NaCl, 0.2 mM phenylmethylsulfonyl fluoride, and the cells were lysed by two cycles of freezing and thawing. Lysates were diluted with SDS-sample buffer, and protein from 100 μl of the original culture was loaded into each lane of a SDS-polyacrylamide gel. Electrophoresis, electrobloading, and reaction with antibodies were as described above.

**RESULTS AND DISCUSSION**

Immune monkey serum was used to screen a *P. falciparum* Camp strain genomic DNA αgt11 expression library with the result that about 150 positive clones were identified (9, 18). Antibodies recognizing expression proteins from individual clones were affinity-purified from the immune serum and used to probe immunoblots of schizont antigens. One clone, a95, was identified as encoding at least a portion of ABRA. Through immunoblot hybridization, and sequence analysis, clone a95 was found to include approximately the 3′-half of the ABRA gene coupled to an interspersed repetitive element (Fig. 1A). The ABRA gene and the repetitive element are not adjacent to each other on the parasite chromosome but rather were joined through an *E. coli* recombination event (19). Overlapping portions of the Camp strain genomic DNA and a FCR3 strain cDNA clone were then isolated by hybridization to the ABRA portion of a95 (Fig. 1A).

Evidence that the cloned DNA encodes ABRA came from reaction of schizont antigens with four antibody preparations: antibodies affinity-purified from total immune serum using the clone a95 expression protein; antibodies affinity-purified using ABRA isolated from parasites; monoclonal antibody 3DS which recognizes ABRA (4); and mouse antisera against a synthetic peptide based on the deduced amino acid sequence of ABRA. All four antibody preparations detected a protein of identical electrophoretic mobility on the schizont antigen immunoblots (Fig. 2).

Monoclonal 3D5 and anti-ABRA antibodies also recognized the β-galactosidase fusion protein expressed by clone a96 (Fig. 3). Sequencing of the 5′ end of the clone a95 insert showed that the ABRA gene was in frame with the vector β-galactosidase gene. The apparent molecular weight of the fusion protein detected on the immunoblots was also consistent with the size of the peptide encoded by the segment of the gene from the EcoRI site to the stop codon. However, the fusion protein is apparently unstable in *E. coli* as it is not detectable by Coomassie Blue staining, and several protein fragments, smaller than the fusion protein, are recognized by the anti-β-galactosidase and anti-101 antibodies (Fig. 3).

A composite gene sequence from the overlapping Camp strain clones is shown in Fig. 4. ABRA is encoded by a single long uninterrupted open reading frame. The ABRA coding region has an A + T content of 75% and an A/T ratio of 1.91 for the mRNAs sense strand. In contrast, the upstream and downstream regions have A + T contents of 88 and 82%, and A/T ratios of 0.79 and 0.87, respectively. Relatively low A + T content and A/T ratios greater than 1 are almost universal features of *P. falciparum* coding regions (20). Two additional open reading frames were detected downstream of the ABRA gene in the opposite orientation (Fig. 1). It is not known if these of open reading frames encodes protein.
likely signal peptidase cleavage site after amino acid 22 (cysteine) (22). ABRA has two regions of tandem peptide repeats (hatched boxes in Fig. 1B). Eight hexapeptide repeats of sequence T\textsubscript{\textsuperscript{N}}\textsubscript{D}ED are encoded by bp 676–819. Near the carboxyl terminus, tandem repeats mostly of sequence KE and KEE are encoded by a polypurine segment of mRNA.

The protein sequence deduced from the long open reading frame has a calculated molecular weight of 86,595 (83,931 for the mature protein), which is considerably less than the apparent molecular weight of 101,000–102,000 estimated from mobility on SDS-polyacrylamide gels (4, 5). True molecular weights for malaria antigens have repeatedly been found to be lower than the apparent molecular weights (15, 23, 24), although the explanation is unknown. There are nine potential N-glycosylation sites of sequence N-X\textsubscript{S}T (X stands for any amino acid) within ABRA, but we have been unable to detect \[^3H\]glucosamine incorporation into ABRA, and, therefore, the high apparent molecular weight is probably not due to glycosylation.

ABRA is a hydrophilic (Fig. 1C) and acidic protein. The calculated pI from the amino acid sequence of the mature protein is 5.6, in agreement with the acidic mobility of ABRA on two-dimensional protein gels.\textsuperscript{7} There are six cysteine residues in ABRA; all are located near the NH\textsubscript{2} terminus (Fig. 1C). Except for the putative signal peptide, there are no protein segments which could obviously anchor ABRA to a membrane. This is consistent with the observation that ABRA is an exoantigen secreted into the parasitophorous vacuolar space (4, 5). The absence of \[^3H\]myristic acid incorporation into ABRA suggests that ABRA is not coupled to lipid after translation, as has been found for gp195, the precursor to several merozoite surface proteins (25).

Searches of version 48 of the GenBank and version 10 of the NBRF data bases revealed no sequences homologous to ABRA. A few proteins from the NBRF data base, notably the yeast heat shock protein 90 (26) and the mammalian neurofilament L protein (27), had short regions rich in lysine and especially glutamic acid, but outside of these regions there was no obvious sequence similarity to ABRA.

ABRA appears to be a conserved malaria antigen. The Camp and FCR3 strains are known to differ significantly in the molecular weights of antigens recognized by growth inhibitory immune serum (28, 29), and in the sequence of a 126,000 molecular weight parasite antigen (4, 18).\textsuperscript{2} In contrast, the sequence of a 1.8-kb FCR3 strain cDNA clone revealed only four differences from the Camp strain genomic sequence (Fig. 4). One was the deletion of an adenine nucleotide within a stretch of 10 adenines (bp 583–592). Deletions or insertions of adenines in runs of adenines have been seen previously in cDNA clones of malaria antigens (30, 31). The other three differences resulted in amino acid changes (N to S at bp 935, E to Q at bp 1078, and D to N at bp 1126). Likewise, the partial FC27 strain (Papua New Guinea) ABRA cDNA sequence (5) is almost identical to the Camp sequence with the only differences being an extra base near the 5' end of the FC27 clone and some rearrangements within the carboxyl terminal repeat region. Additionally, the mobility of ABRA on SDS-polyacrylamide gels is nearly invariant for seven different parasite strains (4, 5).

The most unusual feature of the ABRA gene is the 174-bp polypurine segment (bp 2020–2193). Polypurine-polypyrimidine segments are common in eukaryotes, but previous examples have all been outside of coding regions (32–34). The

---

\textsuperscript{2} J. A. Lyon, J. L. Weber, and J. D. Chulay, unpublished results.
same DNA mechanisms which generate the polypurine-poly- pyrimidine segments in other organisms may have generated (arginine), which would have been found in a random poly-

amino acid sequence must also have been involved since The function of this portion of the

ABRA molecule is un-

Fig. 4. Sequence of ABRA from the Camp strain. The hydrophobic core sequence of the putative signal peptide is underlined. Tandem repeats are bracketed. Base pairs are numbered above the DNA sequence lines; amino acids are numbered to the right of the protein sequence lines. Differences between the FCR3 cDNA sequence (bp 250–2048) and the Camp genomic sequence are shown above the DNA sequence lines (bp 587, 935, 1978, and 1126).
known. An immunological decoy or smokescreen role (35) for this region seems unlikely since human immune serum was found to not react with a synthetic peptide containing KE and KEE repeats (45).

The ABRA gene maps to one of the largest P. falciparum chromosomes (36).

Other P. falciparum blood stage antigens of approximately 100,000 molecular weight (37-41) differ from ABRA by virtue of isoelectric point (37), time of synthesis (38), localization within the infected erythrocyte (39, 40), or DNA sequence (41).

ABRA is found within the parasitophorous vacuole and is released from the parasite by extraction with Triton X-100 or through natural schizont rupture (4, 5). On this basis, Stahl and co-workers (5) concluded that ABRA probably functions as an immunological smokescreen, helping to prevent the host from mounting a strong immune response against antigens which are crucial to parasite survival. However, ABRA has also been located at the schizont and merozoite surface, and the protein is found in clusters of merozoites that have been aggregated with immune serum (4), suggesting that ABRA functions as more than just a smokescreen. ABRA may bind to other integral membrane proteins on the parasite surface (42, 44) and could become a component of a malaria vaccine. Determination of the complete primary structure of ABRA has made rigorous testing of these alternatives possible.

Acknowledgments—We thank Anda Meierovics, Paula May, and Julie Vrana for laboratory assistance and Doreen Luepeke for secretarial help. Gratitude is also extended to Jeff Ravetch for his gift of the FCR3 cDNA library, Bruce Roe and Elaine Mardis for help sequencing the region around the polypurine tract, David Haynes for critical review of the manuscript, and Alan Thomas for assistance with data analysis.

REFERENCES

1. Miller, L. H., Howard, R. J., Carter, R., Good, M. F., Nussenzweig, V., and Nussenzweig, R. S. (1986) Science 234, 1349-1356
2. Chulay, J. D., Aikawa, M., Diggs, C., and Haynes, J. D. (1981) Am. J. Trop. Med. Hyg. 30, 12-19
3. Lyon, J. A., Haynes, J. D., Diggs, C. L., Chulay, J. D., and Pratt-Rossiter, J. M. (1986) J. Immunol. 136, 2252-2258
4. Chulay, J. D., Lyon, J. A., Haynes, J. D., Meierovics, A. I., Atkinson, C. T., and Aikawa, M. (1987) J. Immunol. 139, 2768-2774
5. Stahl, H. D., Bianco, A. E., Crewher, R. F., Anders, R. F., Kyne, A. P., Coppell, R. L., Mitchell, G. F., Kemp, D. J., and Brown, G. V. (1986) Mol. Biol. Med. 3, 351-368
6. Vernes, A., Haynes, J. D., Tachchaisiri, P., Williams, J. L., Dutoit, E., and Diggs, C. L. (1984) Am. J. Trop. Med. Hyg. 33, 197-203
7. Aley, S. B., Sherwood, J. A., and Howard, R. J. (1984) J. Exp. Parasitol. 160, 1585-1590
8. Weber, J. L., Leininger, W. M., and Lyon, J. A. (1986) Nucleic Acids Res. 14, 3311-3323
9. Lyon, J. A., Geller, R. H., Haynes, J. D., Chulay, J. D., and Weber, J. L. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 2989-2993
10. Weber, J. L., and Hockmeyer, W. T. (1985) Mol. Biochem. Parasitol. 15, 305-316
11. Koehan, J., Perkins, M., and Ravetch, J. V. (1986) Cell 44, 689-696
12. Dale, R. M. K., McClure, B. A., and Houchins, J. P. (1985) Plasmid 13, 31-40
13. Johnson-Dow, L., Mardis, E., Heiner, C., and Roe, B. (1987) Biotechniques 5, 1-9
14. Lyon, J. A., and Haynes, J. D. (1986) J. Immunol. 136, 2245-2251
15. Daze, J. B., Williams, J. L., McCutchan, T. F., Weber, J. L., Worth, 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) Science 225, 593-599
16. Goodfied, T. L., Levine, L., and Fasman, G. D. (1964) Science 144, 1344-1346
17. Young, R. A., and Davis, R. W. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 1194-1198
18. Weber, J. L., Lyon, J. A., and Camus, D. (1987) in Molecular Strategies of Parasitic Invasion (Agabian, N., Goodman, H., and Nogueira, N., eds) Vol. 42, pp. 379-388, Alan R. Liss, Inc., New York
19. Weber, J. L. (1988) Mol. Biochem. Parasitol., in press
20. Weber, J. L. (1987) Gene (Amst.) 52, 103-109
21. Simmonds, D., Woollett, G., Bergin-Cartwright, M., Kay, D., and Scaife, J. (1987) EMBO J. 6, 485-491
22. Perlman, D., and Halvorson, H. O. (1983) J. Mol. Biol. 167, 391-409
23. Favaloro, J. M., Coppell, R. L., Corcoran, L. M., Foote, S. J., Brown, G. V., Anders, R. F., and Kemp, D. J. (1986) Nucleic Acids Res. 14, 8265-8277
24. Stahl, H. D., Kemp, D. J., Crewher, P. E., Scanlon, D. B., Woodrow, G., Brown, G. V., Bianco, A. E., Anders, R. F., and Coppell, R. L. (1985) Nucleic Acids Res. 13, 7837-7846
25. Haldar, K., Ferguson, M. A. J., and Cross, G. A. M. (1985) J. Biol. Chem. 260, 4969-4974
26. Farrelly, F. W., and Finkelstein, D. B. (1984) J. Biol. Chem. 259, 5745-5751
27. Geisler, N., Plessmann, U., and Weber, K. (1985) FEBS Lett. 189, 475-478
28. Chulay, J. D., Haynes, J. D., and Diggs, C. L. (1985) Bull. W. H. O. 63, 317-323
29. Vernes, A., Haynes, J. D., Tachchaisiri, P., Williams, J. L., Dutoit, E., and Diggs, C. L. (1984) Am. J. Trop. Med. Hyg. 33, 197-203
30. Triglia, T., Stahl, H. D., Crewher, F. E., Scanlon, D., Brown, G. V., Anders, R. F., and Kemp, D. J. (1987) EMBO J. 6, 1413-1419
31. Pologe, L. G., Pavlovech, A., Shio, H., and Ravetch, J. V. (1988) Proc. Natl. Acad. Sci. U. S. A., in press
32. Levinson, G., and Gutman, G. A. (1987) Mol. Biol. Evol. 4, 203-221
33. Straus, N. A., and Birnboim, H. C. (1976) Biochim. Biophys. Acta 454, 419-423
34. Dodd, J. G., and Straus, N. A. (1982) Biochim. Biophys. Acta 698, 140-148
35. Kemp, D. J., Coppell, R. L., and Anders, R. F. (1987) Annu. Rev. Microbiol. 41, 181-208
36. Weber, J. L. (1988) Exp. Parasitol. 65, 145-158
37. Jendoubi, M., Dubois, P., and Pereira da Silva, L. (1985) J. Immunol. 134, 1941-1945
38. Buranakitjaroen, P., and Newbold, C. I. (1987) Mol. Biochem. Parasitol. 22, 65-77
39. Stanley, H. A., and Reece, R. T. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 6272-6275
40. Gavornik, J. G., Langford, G. J., Crewher, P. E., Saint, R. B., Coppell, R. L., Kemp, D. J., Anders, R. F., and Brown, G. V. (1987) Exp. Parasitol. 63, 58-67
41. Banyal, H. S., and Inselburg, J. (1985) Am. J. Trop. Med. Hyg. 34, 1055-1064
42. Lyon, J. A., Haynes, J. D., Diggs, C. L., Chulay, J. D., Haichris, C. G., and Pratt-Rossitter, J. (1987) J. Immunol. 138, 895-901
43. McBride, J. S., and Heidrich, H.-G. (1987) Mol. Biochem. Parasitol. 23, 71-84