Plasmodium falciparum gene encoding a protein similar to the 78-kDa rat glucose-regulated stress protein

NIRBHAY KUMAR†, CHIANG SYIN‡, RICHARD CARTER§, ISABELLA QUAKYI‡, AND LOUIS H. MILLER‡

*Department of Immunology and Infectious Diseases, †The Johns Hopkins University, School of Hygiene and Public Health, Baltimore MD 21205; and §Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892

Communicated by Richard M. Krause, May 10, 1988

ABSTRACT Genes homologous to heat shock protein 70 have been described in parasitic protozoa. It has been proposed that they may be important to the parasite as it moves from the vertebrate host at 37°C to the insect. We now describe a genomic DNA clone isolated from Plasmodium falciparum that encodes a protein similar to a mammalian heat shock-related protein, the 78-kDa glucose-regulated protein of rat and hamster. The gene is expressed during the erythrocytic stage in both asexual and sexual parasites (RNA blot analysis) and a 72-kDa protein is immunoprecipitated from erythrocytic stage parasites. Importantly, the sequence of the clone is similar to the canonical sequence at the carboxyl termini of glucose-regulated proteins of mammals that determines their localization within endoplasmic reticulum. Since the parasite sequence has only three (Asp-Glu-Leu) of the four carboxy-terminal amino acids, its location and function within the parasite remain to be determined.

Many parasitic protozoa spend part of their life cycle in an invertebrate insect vector and part in a warm-blooded vertebrate host. It is obvious that parasites have evolved mechanisms to adapt to varying physiological environments and shifts in temperature between the two hosts. Van der Ploeg et al. (1) have identified transcripts in Trypanosoma brucei and Leishmania tropica that cross-hybridize with heat shock genes of Drosophila melanogaster and suggested that these genes may play a role in the development of parasites in the invertebrate vector and the vertebrate host. The discovery of these genes in protozoa was not surprising, in that the promoter regions and amino acid sequences of heat shock proteins are conserved throughout evolution (2). In the last 2 years, genes for the major 70-kDa heat shock proteins of Plasmodium falciparum (3–6), T. brucei (7), and Schistosoma mansoni (8) have been cloned and their sequences have been reported. Some heat shock-related genes are expressed constitutively; others are induced by a variety of environmental stresses [e.g., hyperthermia, anoxia, glucose deprivation, and inhibitors of phosphorylation and glycosylation (2, 9)]. A group of these genes encode the 78- and 94-kDa glucose-regulated proteins grp78 and grp94 (10). The grp78 is identical in sequence to the previously identified immunoglobulin-binding protein, Bip (11, 12). Although the heat shock proteins are localized in the nucleus, grp78 and grp94 have signal sequences and remain within the lumen of the endoplasmic reticulum (10, 13). A canonical sequence (Lys-Asp-Glu-Leu) at the carboxyl terminus prevents secretion of grp78 (13). The synthesis of grp78 and grp94 is induced by the presence of malfolded proteins whose transport from the endoplasmic reticulum is blocked (14).

In this paper, we describe a P. falciparum genomic clone that encodes a heat shock-related protein with a high degree of homology to rat and hamster grp78. This homology includes part of the canonical sequence (Asp-Glu-Leu), which may influence its localization and function within the parasite.

MATERIALS AND METHODS

Parasites. Blood-stage asexual and sexual parasites of P. falciparum 7G8 (a clone of Brazilian isolate IMTM22) and 3D7 (a clone of Nijmegen isolate NF54) were obtained from cultures maintained as described (15).

Antibodies. Rabbit antisera were produced by hyperimmunization of rabbits with purified gametes of P. falciparum 7G8 and 3D7 clones. Primary immunization was in complete Freund’s adjuvant and the booster immunizations were in incomplete Freund’s adjuvant.

Construction and Screening of the Library. The library used in these studies was a genomic expression library constructed in the vector Agt11 from mung bean nuclelease-digested genomic DNA from 7G8 parasites (16). Antiserum against 7G8 gametes preabsorbed with lysates of isopropyl-β-D-thiogalactopyranoside-induced Agt11 was used to screen the library as described (17). The positive clones after secondary screening were converted into lysogens using high frequency lysogenic Escherichia coli Y1089. Fusion proteins were purified from the lysates (freeze-thaw) of isopropyl-β-D-thiogalactopyranoside-induced Agt11 using rabbit anti-β-galactosidase antibodies coupled to CNBr-activated Sepharose beads. Fusion proteins still bound to the immunoabsorbent were used to immunize rabbits for the production of antiserum against fusion proteins.

Sequence Analysis of Genomic Clones. EcoRI fragments of Agt11 clones were subcloned into the vector M13mp18 and sequences were determined in both strands by the dideoxy chain-termination method (18). The complete nucleotide sequences of the cloned fragments were obtained using the synthetic oligonucleotides indicated in Fig. 1. The sequence data were analyzed using various sequence analysis programs [e.g., NUCALN, FASTP, and AABANK (Protein Sequence Database of the Protein Identification Resource, National Biomedical Research Foundation) included in the National Institutes of Health Molecular Biology PC-Tools Distribution Package.

Southern and RNA Blot Analyses. DNA from blood-stage asexual parasites (3D7) was digested with various restriction enzymes and mung bean nuclease (19), separated by elec-
trophoresis on 0.8% agarose gels, and transferred to nitrocellulose papers. Total RNA from culture-derived asexual and sexual (gametocytes) parasites (3D7) was isolated by the 4 M guanidinium isothiocyanate/hot phenol extraction method (20), separated on formaldehyde-denatured agarose gels and transferred to nitrocellulose paper. The hybridization probes and conditions are described in the figure legends.

Radiolabeling of Parasites, in Vitro mRNA Translation, Immunoprecipitation and NaDodSO\(_4\)/PAGE Analysis. Asexual and sexual parasite (3D7) proteins were biosynthetically labeled using \(^{35}\)S-methionine (21). Surface proteins in the gametes were labeled by lactoperoxidase-catalyzed radioiodination (21). Total RNA from asexual or sexual parasites was translated in vitro using a rabbit reticulocyte lysate and \(^{35}\)S-methionine. Triton X-100 extracts of labeled parasitem RNA or mRNA-translation products were incubated with 10 \(\mu\)l of various antisera for immunoprecipitation (3 hr at room temperature). Immune complexes were purified using protein A-Sepharose beads and analyzed under nonreducing conditions by 5-15% gradient NaDodSO\(_4\)/PAGE and autoradiography or fluorography (21).

RESULTS

Isolation of Clones. Sera from rabbits immunized with \(P. falciparum\) gametes, the stage of the malaria parasite found in the mosquito midgut, were used to screen a genomic expression library to identify genes encoding gamete surface proteins (22). Here we describe two clones encoding proteins homologous to stress-related proteins of other organisms. One clone is almost identical to the recently described \(P. falciparum\) heat shock protein-like gene (Pfhsps70 (3-5)); the other, T-114, has not been described previously. Both clones (Pfhsps70 and T-114) produced fusion proteins with \(\beta\)-galactosidase, which reacted on immunoblots with anti-\(\beta\)-galactosidase and anti-gamete sera.

Nucleotide and Amino Acid Sequence Analysis of T-114 and Pfhsps70. Subclones of T-114 and Pfhsps70 in the vector M13mp18 were sequenced in both orientations by the dideoxy chain-termination method. Fig. 1 shows the sequences of T-114 and Pfhsps70 (851 nucleotides and 995 nucleotides, respectively). A 63% homology was observed between the nucleotide sequences of T-114 and Pfhsps70 using the NUCALN program. Some gaps were introduced to obtain an optimum match between the two sequences. Both contained single open reading frames, 837 nucleotides for T-114 and 939 nucleotides for Pfhsps70. The sequences of Pfhsps70 and T-114 reported in this paper are not the full-length sequences. Both clones represent sequences of proteins starting in the middle and extending to the carboxy terminus. The deduced amino acid sequence of clone T-114 was compared with the sequences in the AABANK (protein sequence database). Clone T-114 showed significant sequence similarity (53%) with Drosophila 70-kDa heat shock protein hsp70 (residues 352-643) and Pfhsps70. In addition, it showed 63% similarity to the carboxy-terminal residues 377-654 of the grp78 of rat and hamster (11, 23) (Fig. 2). Comparison of carboxy-terminal sequences of T-114 and various other glucose-regulated proteins also suggested that T-114 is more similar to grp78 than to other such proteins, although all the glucose-regulated proteins have the same last four carboxy-terminal amino acids (Fig. 3a). Twenty-eight out of 38 carboxy-terminal amino acids, including the last three (Asp-...).
Glu-Leu), were identical between the T-114-encoded protein and rat grp78 (Fig. 3a).

The sequence of Phsp70 (deduced amino acid sequence of the clone in this paper) is almost identical to the recently published sequences of CDNA clones of *P. falciparum* (3–5), with two exceptions: replacement of alanine for glycine and of arginine for alanine at the positions indicated by the asterisks in Fig. 2. As noted previously, Phsp70 has seven repeats near the carboxyl terminus [five repeat sequences of Gly-Gly-Met-Pro and two containing insertions within the repeat sequence (Fig. 3b)]. Phsp70 in the FCR-3 parasite also contained an additional repeat sequence (Ser-Gly-Met-Pro) (5). The Gly-Gly-Met-Pro repeats observed in Phsp70 are not unique to *Plasmodium* parasites, as similar repeat sequences have been reported in *T. brucei* (7). Even *Xenopus* and yeast hsp70 have been found to contain tandem tetrapeptide sequences in this region (Gly-Gly-Val-Val-Gly-Val-Val-Gly-Gly-Met-Pro and Gly-Gly-Ala-Pro-Gly-Gly-Ala-Ala-Gly-Gly-Pro-Gly-Gly-Ala-Pro, respectively) (Fig. 3b).

**Southern and RNA Blot Analysis.** Nick-translated plasmid DNAs (replicative form) from M13 subclones of T-114 and Phsp70 were used as probes in Southern and RNA blot analyses. In the Southern blot analysis, the T-114 insert hybridized to a single DNA band in DNA digests that did not restrict within the clones (Fig. 4D, lanes 4–7). This pattern is different from that of Phsp70 where multiple bands are detected (Fig. 4E), probably representing multiple genes for hsp70. DNA from both clones hybridized to a broad RNA band having an average size of 2.8 kilobases (kb) from both asexual (lanes 1) and sexual (lanes 2) parasites on RNA blots (Fig. 4A and B). Phsp70 also hybridized faintly with an RNA band of slightly smaller size. To confirm that the hybridization was specific for T-114 and Phsp70 mRNA transcripts, we used as probes on RNA blots oligonucleotides corresponding to the unique sequences toward the 3' end of each coding sequence. As shown in Fig. 4C, the oligonucleotide specific to T-114 hybridized to an RNA species of 2.9 kb, whereas the oligonucleotide specific for Phsp70 hybridized to bands of 2.8 and 3.3 kb.

**Identification of the T-114-Encoded Polypeptide in Parasites.** Rabbit antisera against purified fusion proteins were used to immunoprecipitate proteins from Triton X-100 extracts of labeled parasites and mRNA-dependent *in vitro* translation products. Antiserum against the fusion protein of clone T-114 immunoprecipitated a protein (sometimes seen as a doublet) of apparent molecular mass of 72 kDa from [35S]methionine-labeled gametocytes (Fig. 5A, lane 4). This protein is consistently of lower molecular mass than the protein immunoprecipitated by antiserum to the fusion protein of Phsp70, which is 75 kDa (Fig. 5A, lane 3). In contrast to these proteins synthesized *in vivo*, the 72- and 75-kDa proteins from *in vitro* translation of mRNA are of similar size. The higher molecular mass of the T-114-encoded protein observed in *in vitro* translation of mRNA probably reflects *in vivo* processing (Fig. 5B, lanes 3 and 4).

The antiserum to the T-114 fusion protein also immunoprecipitated proteins of 230 and 200 kDa from metabolically labeled gametocytes (Fig. 5A, lane 4) and a 230-kDa protein from the surface radiolabeled gametocytes (Fig. 5C, lane 4). The 230-kDa surface protein is a target of antibodies that block infectivity of parasites to mosquitoes and is a candidate antigen for a malaria transmission-blocking vaccine (32). The T-114 clone cannot, however, encode the 230-kDa protein, since the T-114 mRNA is only 2.9 kb, which is too small to encode a protein of this size. We have also produced antisera in rabbits against the peptide corresponding to the last 12 amino acid residues of the T-114 open reading frame (see Fig. 2). This antisera immunoprecipitated the 72-kDa protein as did the antiserum against the fusion protein of clone T-114 (data not shown). The 230-kDa protein was not immunoprecipitated by the anti-peptide serum.

The antisera against the two fusion proteins were also used to detect proteins by immunoblotting. The antiserum against T-114 reacted with the 72-kDa protein; however, no reactive bands were found in the region of 230 kDa (Fig. 5E, lane 4). The reactivity of the antiserum against the fusion protein of Phsp70 to proteins in asexual and sexual parasites is shown in Fig. 5 D and E (lanes 3). Both sera gave positive reactions with antigens in methanol-fixed erythrocytic asexual and
FIG. 4. RNA and Southern blot analyses. (A–C) RNA blot. 32P-labeled DNAs from Pfhs70 (A) and T-114 (B) were hybridized with RNA from asexual (lanes 1) and gametocytes (lanes 2). Hybridization was done at 55°C [4 x SSC (1 x SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7)/0.1% NaDodSO4] and blots were washed at 55°C with 0.2 x SSC/0.1% NaDodSO4. Positions of rRNAs are identified with dashes on the left of A. (C) End-labeled oligonucleotide probes specific for Pfhs70 (lane 1) and T-114 (lane 2) were hybridized with RNA from asexual parasites. Hybridization was done at 42°C (4 x SSC/0.1% NaDodSO4) and blots were washed using 2 x SSC/0.1% NaDodSO4 at 42°C. (D and E) Southern blot. DNAs (1 μg per lane) were digested with Mbo II (lane 1), Kpn I (lane 2), HindIII (lane 3), HindIII (lane 4), EcoRI (lane 5), Dra I (lane 6), or mung bean nuclease in the presence of 40% formamide (lane 7) and then hybridized with 32P-labeled DNA from T-114 (D) and Pfhs70 (E). Hybridization was done at 65°C and blots were washed with 0.2 x SSC/0.1% NaDodSO4 at 50°C.

sexual parasites as well as with gametes in indirect immunofluorescence assays. In addition, the antiserum against the T-114 fusion protein showed a positive reaction on the surface of intact gametes (data not shown).

DISCUSSION

We have described the isolation of two genomic clones: one that is identical in sequence (with a few substitutions) to the

FIG. 5. Immunoprecipitation and immunoblotting analysis. (A–C) Immunoprecipitation. Triton X-100 extracts of [35S]methionine-labeled gametocytes (A), in vitro translation products of mRNA from asexual parasites (B), and surface radioiodinated gametes (C) were immunoprecipitated with various antibodies: normal rabbit serum (lane 1), rabbit anti-gamete serum (lane 2), rabbit antiserum against the Pfhs70 fusion protein (lane 3), and rabbit serum against the fusion protein of T-114 (lane 4). (D and E) Immunoblotting. Nitrocellulose strips containing asexual (D) and gametocyte (E) antigens were treated with preimmune (lanes 1 and 2) and immune (lanes 3 and 4) sera from rabbits immunized with the fusion proteins of Pfhs70 (lanes 1 and 3) and T-114 (lanes 2 and 4). Lane 5: the strip was treated with rabbit antiserum against the 230-kDa protein (31). 125I-labeled protein A was used to detect binding of antibodies to antigens.
recently published PfP70 (3–5) and T-114 sequence that encodes a protein with a high degree of homology to the carboxyl-terminal region of the 78-kDa glucose-regulated protein of rat and hamster (grp78) (11, 23). Clone T-114 contains internal homology to grp78 and hsp70 and homology to the carboxyl-terminal region of grp78. Based on this sequence homology, we suggest that the T-114-encoded protein may be a glucose-regulated protein. Confirmation of its identity with grp78-like proteins, however, must await physiological data on responses to glucose deprivation and other known stimuli, sequence data on the amino terminus, where the gene should have a signal for secretion into the endoplasmic reticulum, and ultrastructural localization in the lumen of the endoplasmic reticulum.

One interesting difference between the deduced amino acid sequence of T-114 and that of grp78 is the absence of a lysine in the malariial protein (which has the sequence Ser-Asp-Glu-Leu) instead of Lys-Asp-Glu-Leu, thought to be the canonical sequence for localization of grp78-like proteins in the lumen of the endoplasmic reticulum. In a study using various deletion and insertion plasmids, Munro and Pelham (13) have shown that the presence of the Lys-Asp-Glu-Leu sequence at the carboxyl terminus of proteins such as grp78, grp94, and protein disulfide isomerase is necessary for their retention in the lumen of the endoplasmic reticulum. We do not know whether lack of the positively charged amino acid lysine in parasites would have an effect on the compartmentalization in parasites or the parasite uses a different signal than mammalian cells.

The antiseraum against the T-114 fusion protein immunoprecipitated not only the 72-kDa protein but also two other parasite proteins. One of these is a 230-kDa gamete surface antigen, a candidate immunogen for malaria transmission-blocking immunity (32). T-114 does not encode the 230-kDa protein because it hybridizes to an mRNA of only 2.9 kb, which is too small to encode this protein. It is not clear from our data whether the 230-kDa gamete surface protein shares a cross-reacting epitope with the T-114-encoded protein or whether the 230-kDa protein is immunoprecipitated with the 72-kDa protein.

In general, the exact functions of these heat shock-related proteins have not been established. For the Plasmodium family of these stress proteins we also do not know whether they are expressed constitutively or are inducible by various physiological and developmental stimuli. In either case, they may play a role in the adaptation of parasites during their complex life cycle. Characterization of expression of these genes both at the transcriptional and at the translational level in various stages of the parasite in the vertebrate host and in the insect vector will help in understanding their roles in the normal growth and development of the parasite.

We thank Dr. Thomas McCutchan for making available to us the genomic library, Dr. Teresa M. Y. Liu for synthesizing oligonucleotides, Dr. Paul Englund for critical comments on the manuscript, and Mrs. H. Bongers for editorial assistance. The work was supported by National Institutes of Health Grant AI-24704-02, the United Nations Development Programme/World Bank/World Health Organization Special Programme for Research and Training in Tropical Diseases, and the John D. & Catherine T. MacArthur Foundation.