Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 11889–11894, October 1996
Medical Sciences

Comparison of circumsporozoite proteins from avian and mammalian malarias: Biological and phylogenetic implications

(plasmodium/apicomplexa/evolution/sporozoite invasion/cell adhesion motif)

THOMAS F. MCCUTCHAN*†, JESSICA C. KISSINGER*, MUSA G. TOURAY*‡, M. JOHN ROGERS*, JUN LI*, MARGERY SULLIVAN*, ERIKA M. BRAGA§, ANTONIANA U. KRETTLI§, AND LOUIS H. MILLER*

*Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892-0425; and §Centro de Pesquisas René Rachou, FIO Cruz and Federal University of Minas Gerais, Belo Horizonte 30 190–002, Minas Gerais, Brazil

Contributed by Louis H. Miller, July 31, 1996

ABSTRACT **The circumsporozoite (CS) protein of malaria parasites (*Plasmodium*) covers the surface of sporozoites that invade hepatocytes in mammalian hosts and macrophages in avian hosts. CS genes have been characterized from many *Plasmodium* that infect mammals; two domains of the corresponding proteins, identified initially by their conservation (region I and region II), have been implicated in binding to hepatocytes. The CS gene from the avian parasite *Plasmodium gallinaceum* was characterized to compare these functional domains to those of mammalian *Plasmodium* and for the study of *Plasmodium* evolution. The *P. gallinaceum* protein has the characteristics of CS proteins, including a secretory signal sequence, central repeat region, regions of charged amino acids, and an anchor sequence. Comparison with CS signal sequences reveals four distinct groupings, with *P. gallinaceum* most closely related to the human malaria *Plasmodium falciparum*. The 5-amino acid sequence designated region I, which is identical in all mammalian CS and implicated in hepatocyte invasion, is different in the avian protein. The *P. gallinaceum* repeat region consists of 9-amino acid repeats with the consensus sequence QP(A/V)GGNGG(A/V). The conserved motif designated region II-plus, which is associated with targeting the invasion of liver cells, is also conserved in the avian protein. Phylogenetic analysis of the aligned *Plasmodium* CS sequences yields a tree with a topology similar to the one obtained using sequence data from the small subunit rRNA gene. The phylogeny using the CS gene supports the proposal that the human malaria *P. falciparum* is significantly more related to avian parasites than to other parasites infecting mammals, although the biology of sporozoite invasion is different between the avian and mammalian species.**

Comparisons of homologous proteins from distantly related *Plasmodium* species have suggested domains retained for function. Two examples are conserved regions of the circumsporozoite (CS) protein of sporozoites (1) and the conserved regions of a family of erythrocyte binding proteins on merozoites (2). Sporozoites are the infective form of malaria parasites that are inoculated by mosquitoes into the vertebrate host. The merozoite is the parasite stage that invades erythrocytes. A comparison of CS protein sequences from *Plasmodium falciparum* and the distantly related parasites, *Plasmodium knowlesi* and *Plasmodium vivax,* identified domains (regions I and II) that were shown to be conserved (1) and there is evidence suggesting that these domains function in sporozoite binding to hepatocytes (3, 4). A similar approach identified domains on a family of merozoite proteins that were demonstrated to bind erythrocytes (5, 6).

Although the processes whereby the sporozoite develops and infects the invertebrate and vertebrate hosts are similar among *Plasmodium* species, there are a few fundamental differences (7). For example, sporozoites of avian malaria parasites develop predominantly in culicine mosquitoes, while the primate malarias use anopheline vectors. Sporozoites of avian parasites infect macrophages; sporozoites of primate malarias infect hepatocytes. These biological differences may be reflected in differences in the functional domains of the CS protein, an important ligand on the sporozoite surface.

The sequence of the CS gene has been determined for a number of species that infect mammals but no sequence has been determined for an avian malaria parasite. Herein we compare the sequence of the CS protein gene of an avian parasite, *Plasmodium gallinaceum*, with corresponding sequences from mammalian parasites. Analysis of the CS protein genes supports the evolutionary relationship between avian parasites and *P. falciparum*, despite similarity of the biology of *P. falciparum* sporozoites to other mammalian *Plasmodium* species. In addition we focus on comparisons of regions I and II that have been implicated in hepatocyte invasion for mammalian *Plasmodium* species.

MATERIALS AND METHODS

Parasites. *P. gallinaceum* was maintained in chickens and infected *Aedes aegypti* maintained at 28°C and 80% relative humidity. The infection of the mosquitoes was checked at day 5 for the presence of oocysts and at day 10 for the presence of sporozoites in the salivary gland.

Preparation of RNA and DNA. RNA was isolated from 10^7 sporozoites purified from the salivary glands from approximately 1000 *P. gallinaceum*-infected *Ae. aegypti* mosquitoes as described (8). Total RNA was purified from the sporozoites by organic extraction (9). DNA was purified from 50 ml of *P. gallinaceum*-infected chicken blood that had a parasitemia of 70%. Total nucleic acids were isolated by organic extraction of the lysate of *P. gallinaceum*-infected erythrocytes produced by SDS/proteinase K treatment (9). *P. gallinaceum* DNA was separated from host DNA and other nucleic acids by Hoechst dye/CsCl gradient centrifugation as described (9).

Competitive Indirect Immunofluorescence (IIF). The synthetic peptide 3893 (GGVQPAGGNGGVQPAGGNGGVQPAGGN-amide) corresponding to a portion of the repeat region of the CS of *P. gallinaceum* was used at different concentrations (1 μg/ml, 10 μg/ml, 25 μg/ml, 50 μg/ml, and 100 μg/ml) to compete for binding of the mAbs N2H1D5 and N5G^{3}H6 to *P. gallinaceum*

Abbreviation: CS, circumsporozoite.

Data deposition: The sequence reported in this paper has been deposited in the GenBank data base (accession no. U65959).

†To whom reprint requests should be addressed at: Growth and Development Section, Laboratory of Parasitic Diseases Building 4, Room B1-28, 9000 Rockville Pike, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892-0425. e-mail: mcutchan@helix.nih.gov.

‡Present address: Department of Rheumatology, University Hospital of Lausanne, 1011 Lausanne, Switzerland.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

sporozoites (10). A nonrelated peptide (NANPNVDPNANP) corresponding to the repeat of the CS of *P. falciparum* was used as a control in the competition experiments at a concentration of 100 μg/ml. Dilutions of the mAbs were incubated with the different concentrations of peptides at various dilutions for 3 h at 37°C. The mixture was incubated with air dried sporozoites on slides for 30 min at 37°C. Slides were washed and incubated with fluorescein-conjugated anti-mouse IgG (γ-chain specific, Sigma) for 30 min at 37°C. The washed slides were examined for immunofluorescence. Each dilution and antigen concentration was coded and read blind.

Isolation and Sequence Analysis of the CS Protein Gene. cDNA was made from RNA isolated from sporozoites using the oligo(dT) priming as described (11). A recombinant cDNA library was made in λ-ZAP II (Stratagene) as described by the manufacturer. The mAb-specific for the *P. gallinaceum* circumsporozoite protein N5G^{3}H6 (10) was used to screen the cDNA library for phage containing the CS protein gene (12).

A genomic library in λ ZAP Express (Stratagene) was prepared as described by the manufacturer, carrying fragments of DNA obtained from a partial digest of *P. gallinaceum* DNA with the restriction endonuclease *Sau*3AI. Recombinant phage that contained the CS protein gene were detected by hybridization to a radiolabeled oligonucleotide, Pg1022 (5′-CTGGAGGWAATGGTGGTGGT-3′), that is homologous to the repeated sequence found in all cDNA clones detected by the mAb N5G^{3}H6. Excision of the recombinants as phagemid pBK-CMV was as described by the manufacturer. Digestion of phagemid DNA with mung bean nuclease in the presence of 40% formamide as described (13) yielded a single 1.25-kb fragment containing the complete CS protein gene. This fragment was subcloned in the *Srf*I site of pCR-Script SK(+) (Stratagene) as described by the manufacturer and sequenced by the dideoxynucleotide chain-termination method with oligonucleotides 250–300 nt apart as sequencing primers. The sequence was confirmed on both strands, and oligonucleotides made to central variant repeats confirmed the number and sequence of the repeats.

Phylogenetic Analysis. Circumsporozoite protein sequences from *P. gallinaceum* and 17 other *Plasmodium* species were aligned with CLUSTAL W (14) and the alignment was then refined by eye. Only the N-terminal 74 residues and the C-terminal 87 residues were unambiguously aligned and included in subsequent analyses. The data set was analyzed with PAUP 3.1.1 (D.L. Swofford, Champaign, IL) with 100 replicates of a heuristic search using random addition. The data set was then subjected to bootstrap analysis 100 times, with each replicate consisting of 10 heuristic searches using random addition. The corresponding nucleotide data set was constructed and analyzed the same way.

RESULTS

Identification of the Immunodominant Repeated Epitope. The CS protein of *P. gallinaceum,* like the CS protein of other malaria parasites, includes an immunodominant repetitive epitope recognized by anti-sporozoite mAbs (10). All the anti-sporozoite mAbs recognize a protein on the sporozoite surface and recognize a repetitive epitope as determined by a two-site one-antibody assay (10). Using the anti-CS mAb (N5G^{3}H6), recombinant phage from a *P. gallinaceum* sporozoite-specific cDNA library were isolated that express this epitope. Sequence analysis of the DNA inserts indicated that all contained a repeating sequence 27 nt long with some sequence variation (data not shown). These sequences correspond to a portion of the central repeat sequences obtained from a genomic DNA clone described below (Fig. 1).

The mAb N5G^{3}H6 reacted with sporozoites in a two-site assay (10). Therefore, it should bind to the central repeat region expressed by the cloned gene. At a concentration 50-fold above the endpoint in the IIF test, the mAB was preincubated with of a peptide at 100 μg/ml derived from the repeat sequence of the CS protein of *P. gallinaceum*. After incubation, the antibody did not bind to sporozoites as detected by the IIF test. A control peptide derived from the repeat of the CS protein of *P. falciparum* did not inhibit IIF at the same concentration. The finding that a peptide derived from the cloned repeat sequence blocks the binding of an anti-CS mAb to *P. gallinaceum* sporozoites is consistent with fact that the repeat is from a portion of the CS gene.

Isolation of the Complete CS Protein Gene from *P. gallinaceum*. The CS protein gene was isolated from a genomic library of *P. gallinaceum* DNA. Genomic DNA used for the construction of the *P. gallinaceum* library was purified from total DNA isolated from *P. gallinaceum*-infected erythrocytes from a chicken. Most of this material was chicken DNA because avian red blood cells are nucleated. Parasite DNA was separated from the host DNA by CsCl centrifugation in the presence of Hoechst dye as described (9). DNA fragments, averaging approximately 10 kb in size, that resulted from partial digestion of parasite DNA with the restriction enzyme *Sau*3AI were used to construct a genomic library. The library was screened for CS protein gene-related fragments, using a radiolabeled probe homologous to the repeated region described above. Approximately 20,000 phage were screened yielding phage isolates that contained sequence homologous to the probe. The insert of one of these was excised as a clone in the kanamycin-resistant pBK-CMV phagemid. The phagemid contained an insert of approximately 5 kb. The CS gene was then cleaved from the phagemid using mung bean nuclease with conditions that have been shown to excise *Plasmodium* genes as complete units from genomic DNA (13). This yielded a fragment approximately 1.2 kb long that was subsequently cloned and sequenced in the pCR-Script vector (Fig. 1).

The Sequence of the *P. gallinaceum* CS Protein Gene. The sequence of the cloned *P. gallinaceum* gene is similar to the corresponding CS protein genes from *Plasmodium* that infect mammals (1). The ATG initiation codon of the mung bean fragment is at position 78 and terminates with a TAA codon at position 1245. The first 20 amino acids of the N-terminal end of the protein constitutes a probable secretory signal peptide sequence, with predicted cleavage between amino acids 20 and 21 (15). The alignment (Fig. 2) shows that the secretory signal sequences can be grouped into four categories (avian with *P. falciparum* A/F; *Plasmodium malariae* M; primate malarias P; rodent malarias R). The *P. gallinaceum* and *P. falciparum* sequences appear to constitute a distinct grouping in which 18 of the first 23 amino acids of the *P. gallinaceum* sequence are identical to those found in the *P. falciparum* sequence.

The sequence between the secretory signal sequence and the repeated region is highly charged, as is the case in all CS proteins to date. The region in *P. gallinaceum* is somewhat less charged than the corresponding region of CS proteins from malaria parasites of mammals and contains a high number of asparagine residues. The region immediately N-terminal to the repeats has been designated region I (1) and is identical in all malaria parasites of mammals so far described (Fig. 3). There is no significant identity seen when a corresponding region of the *P. gallinaceum* sequence is included in the alignment. To maximize the extent of possible identity alignment we have included the first repeat unit in the sequence (Fig. 2). As shown in Fig. 3, the core of the region I area, amino acids KLKQP, is highly conserved in all *Plasmodium* species aligned, with the exception of *P. gallinaceum*. The *P. gallinaceum* CS sequence contains two amino acid changes within the core, a unique insertion immediately adjacent to the core, and numerous changes in the flanking regions.

The repeated region encodes 16 repeats of a 9-amino acid sequence whose consensus is QP(A,V)GGNGG(A,V). Four of the central repeats vary considerably from the consensus and may be immunologically distinguishable from the other repeats (Fig. 1). It is known that different mAbs that bind to the repeated sequences vary considerably in regard to their effect

FIG. 1. Nucleotide sequence of the *P. gallinaceum* fragment obtained by mung bean nuclease cleavage. The predicted open reading frame of the CS protein is shown. The predicted secretory signal sequence is shown with an open dashed boxed (15). Regions showing similarity to region I and region II-plus as discussed in text are indicated by shaded overlines, with conserved sequences in shaded boxes. The repeat regions are underlined, with variant central repeats indicated by open boxes. Two fragment sizes are obtained upon PCR amplification of the repeated region. An exact insertion of 10 additional repeats, after unit 12, represents the only sequence difference between the two.

in blocking invasion of macrophages (16), but we do not know that the discriminatory effect of these antibodies is the result of a distinction based upon primary sequence variation. Also of interest is the fact that the consensus repeat sequence found in this avian parasite is very similar to that found in the *Plasmodium cynomolgi* complex, a monkey malaria. The consensus core of the *P. cynomolgi* and *P. vivax* repeats (17) is present in *P. gallinaceum* CS not only at the amino acid level but also in 11 of 12 nucleotide positions.

After the repeat region, there is another highly charged region similar to that seen in other CS proteins. Within this region is the 18-amino acid sequence designated region II-plus (18). This region was identified from homology between the *P. falciparum* and *P. knowlesi* (1) and later designated region II-plus as the sequences of more *Plasmodium* species became available and from inhibition of binding studies (18). Region II-plus has been found in all *Plasmodium* CS genes sequenced to date (18) and also conserved in other mammalian proteins from a variety of sources, including two additional cysteines C-terminal to region II-plus (19). It has been found to be associated with cell–cell adhesion in general (20, 21) and, more specifically as it relates to malaria, with the targeting of sporozoites to hepatocytes (3, 18). The sequence common to the malaria parasites of mammals (EWXXCXVTCGXGV/IXXRXK/R; position 15 has an aliphatic side chain) has been shown to be involved in liver invasion. The only difference in the *P. gallinaceum* CS to the region II-plus consensus sequence is the conservative arginine to lysine substitution at position 16. The C-terminal end of the CS is composed of hydrophobic amino acids, consistent with other CS proteins (1) with this probably being a sequence for attachment of glycosylphosphotidylinositol (GPI).

Phylogenetic Analysis. Only the N- and C-terminal regions of the *P. gallinaceum* CS protein were unambiguously aligned with other *Plasmodium* CS proteins (Fig. 2). Parsimony analysis of the aligned regions of the CS peptide sequence produced five equally parsimonious trees of length 330 and a consistency index (C.I.) of 0.873. The only differences between the trees are the relationships at the extremities of the *P. vivax* and *P. falciparum* branches. One of the trees is shown in Fig. 4. Analysis of the nucleotide data set yielded six equally parsimonius trees of the same overall topology that differ as described above. Using both data sets, *P. gallinaceum* groups with the *P. falciparum*/*Plasmodium reichenowi* lineage with significant bootstrap support. The relationships of the other

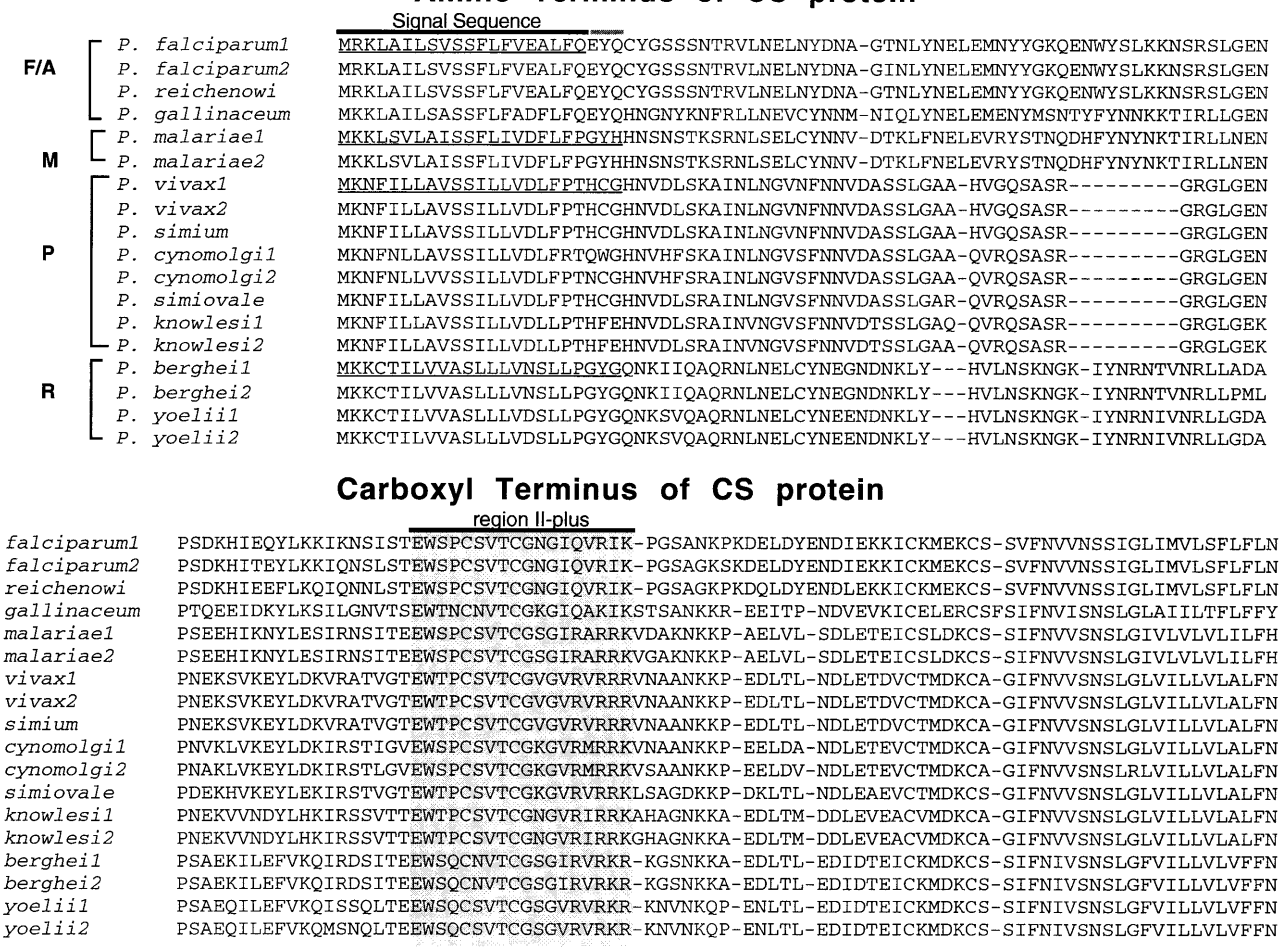

FIG. 2. Alignment of the N- and C-terminal regions of the circumsporozoite protein from different *Plasmodium* species. This alignment served as the data set in the analyses used to generate the phylogram in Fig. 4. The sequences are presented in the order in which they appear in Fig. 4 for ease of comparison. The putative signal sequence is underlined in the groups divided into *P. falciparum*/avian *Plasmodium* (F/A), *P. malariae* (M), primate *Plasmodium* (P), and rodent *Plasmodium* (R). The primate malarias include those *Plasmodium* specise that evolved from Old World monkeys to higher primates and do not include the human malarias *P. falciparum* and *P. malariae*.

major *Plasmodium* lineages are consistent with a previous analysis on the CS protein (22) and lend support to the strength of the phylogenetic signal contained in the CS molecule. Both analyses have obtained a similar result, although the study of Escalante *et al.* (22) used a different phylogenetic inference method. The results of our analyses are comparable with those obtained from an analysis of small subunit rRNA sequences (23) that showed the relationship of *P. falciparum* to avian *Plasmodium*. A feature of the CS protein of importance is its immunogenicity. The most rapidly evolving and immunogenic

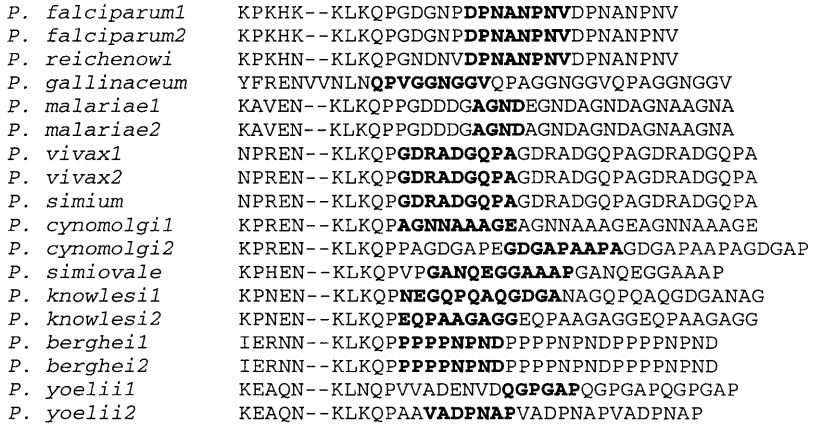

FIG. 3. Comparison of the region I containing amino acids immediately upstream of the first repeat sequence. The region I core is marked by asterisks and the first repeat is indicated in boldface type.

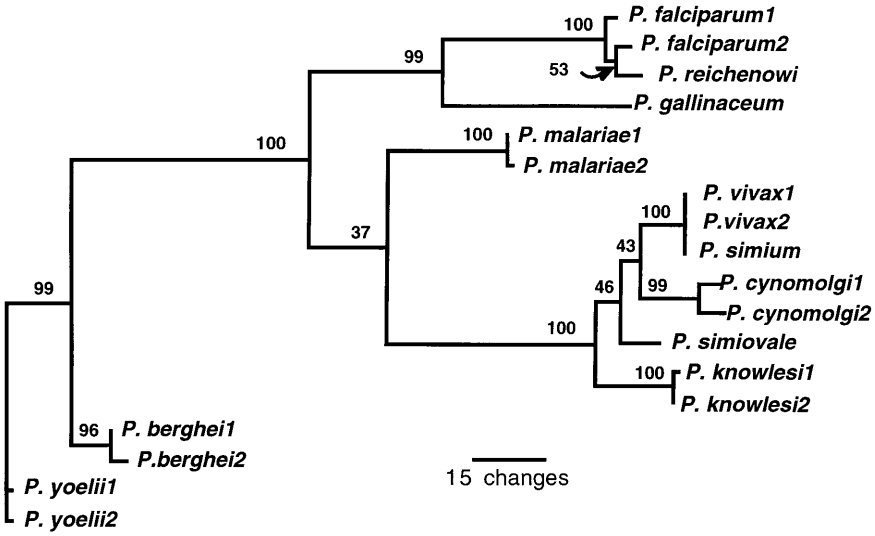

FIG. 4. Phylogram of circumsporozoite protein relationships. The tree was derived with parsimony analysis on the amino acid sequences shown in Fig. 2 (the repeat regions were excluded). Numbers shown above the branches are bootstrap percentages based on 100 replicates. GenBank accession numbers for the CS sequences used in this anaylsis are: *P. falciparum,* K02194 and M83164; *Plasmodium reichenowi,* M60972; *P. gallinaceum,* U65959; *P. malariae,* J03992 and U09766; *P. vivax,* M11925 and M34697; *Plasmodium simium,* L05068; *P. cynomolgi,* M15103 and M15104; *Plasmodium simiovale,* U09765; *P. knowlesi,* K00822 and M11031; *Plasmodium berghei*, M14135 and M28887; *Plasmodium yoelii*, J02695 and M58295.

portions of the CS protein are the repeat regions, with polymorphism in these regions (24, 25). These regions cannot be aligned and were excluded from the analysis. However, the immunogenic nature of this protein raises potential questions about its usefulness as a phylogenetic marker. In this light, it is interesting to note that parasites that infect the same hosts, in this case humans, and presumably face the same immunological pressures, do not group together. Instead, when repeat regions are removed, *P. malariae*, *P. vivax*, and *P. falciparum* form distinct lineages (Fig. 4).

DISCUSSION

The report of an avian CS protein gene sheds light on some basic differences between the biology of avian and mammalian malaria sporozoites. In review, all *Plasmodium* sporozoites develop in the oocyst that forms on the outside of the mosquito midgut. Upon maturation and rupture of the oocyst wall, the sporozoites move through the mosquito hemocoel and enter the salivary gland. From the salivary gland, the sporozoites pass into the vertebrate host through the bite of an infected mosquito. There are, however, some striking differences in the manner sporozoites develop in the mammalian and avian host. At least one of these is associated with the function of the CS protein. In mammalian *Plasmodium* parasites, primary infection of the host and initial development of the parasite occurs in hepatocytes. Avian malarias are different in that, after introduction into the vertebrate host, the initial development occurs in macrophages. These cells also support sporozoite development *in vitro* for the formation of exoerythrocytic merozoites that are infectious (26).

It was previously anticipated that the conservation of sequence motifs among the CS genes of mammalian parasites may be an indicator of regions important to sporozoite function (1). Including an avian *Plasmodium* CS protein sequence in the comparison of *Plasmodium* CS sequences may indicate differences corresponding to biological differences between avian and mammalian parasites. The sequence comparison of all mammalian parasite CS proteins revealed only two short regions of similarity and these are referred to as region I and region II-plus. Region II-plus is a common motif seen in the proteins of a wide range of organisms and has been shown to be associated with cell-cell interactions (18–21). A number of studies have convincingly shown that region II-plus in sporozoites is involved with targeting the liver in the vertebrate host (3, 18). The remaining question is whether or not it is the sole factor involved in hepatocyte targeting. The involvement of region I in liver invasion has also been suggested with the caveat that protein processing near region I may potentiate the process. In light of the above, the addition of an avian *Plasmodium* CS protein sequence to the alignment might have been anticipated to have shown a missing or significantly altered region II-plus region. This is not the case; region II-plus remains intact, while region I is altered. This suggests that region II-plus is essential to targeting the invasion of both avian and mammalian sporozoites and the process is mechanistically similar in both. This is of interest to those studying the developmental process of the parasite as well as those involved in testing vaccines in animal models. Region II-plus may still be the only factor involved, and differences between the primary site of invasion could simply reflect physiological differences between birds and mammals. The fact that there is no recognizable region I in this avian parasite should renew interest in the function of region I and its possible involvement in the specificity of invasion. It remains to be established whether recombinant avian CS protein targets avian macrophages, analogous to recombinant mammalian CS protein targeting hepatocytes.

The phylogenetic association of malaria parasites has been determined both on the basis of taxonomic characters and in terms of molecular phylogeny. The assessment of phylogenetic relationships of *Plasmodium* species has been with the assumption of coevolution of host and parasite. This premise would imply that parasites that infect humans are more likely to be phylogenetically related to each other than to parasites from monkeys, birds, or lizards. It has long been noted that *P. falciparum* has a striking resemblance to avian *Plasmodium*. Early discussions focused on the fact that both avian *Plasmodium* and *P. falciparum* has crescentic gametocytes (27–29). Later, ultrastructural analyses of *Plasmodium* mitochondria demonstrated that avian *Plasmodium* and *P. falciparum* had cristate mitochondria while the other *Plasmodium* species did not (30). Support for the origin of *P. falciparum* from avian *Plasmodium* was also provided by studies of the nucleotide composition of *Plasmodium* genomes (31). The genomic composition studies revealed that malaria parasites fall into three separate and distinct categories, with the DNA of *P. falciparum* and several avian parasites falling into the group defined by a 17% G + C content. Recently, phylogenetic

analysis of molecular sequence data also supports the association of *P. falciparum* and avian malaria parasites. Waters *et al.* (24) compared small subunit rRNA sequences from nine *Plasmodium* species, determined direct allelic relationships on the basis of function and time of expression, and determined from this information that avian parasites formed a monophyletic group with *P. falciparum.* This suggests either a lateral transfer of parasites between bird and human hosts or some common progenitor of both that was not shared by other human malaria parasites. This phylogenetic analysis of rRNA sequences proved to be controversial and led to much discussion, although the historical context of this argument was perhaps ignored (32–35). For example, using these data and including two unpublished GenBank entries of putative reptilian *Plasmodium* species, Escalante and Ayala (36) performed a similar analysis and showed that the addition of the reptilian sequences to the data set resulted in two equally favored trees, one that supported the avian malaria—*P. falciparum* association and one that did not. The importance of this conclusion awaits proof that the sequences isolated by polymerase chain reaction of lizard DNA originated from their infection with *Plasmodium mexicanum* and some evidence that the genes being compared were orthologous with the previously compared genes. This last point is crucial since *Plasmodium* species have multiple, divergent, differentially expressed rRNA genes (37). Currently, no reptilian CS gene sequence is available for inclusion in this study, but efforts are underway to clone this gene.

To elucidate the origins of *P. falciparum* and its relationship to avian *Plasmodium*, we have examined a second molecule, the CS protein. Escalante and Ayala (22) have previously reported a phylogenetic analysis of the mammalian CS protein genes. At the time of their study, the sequence from an avian *Plasmodium* was not available. Herein we have included the sequence of *P. gallinaceum* in a similar phylogenetic analysis, using both an alignment kindly provided by Escalante and Aylala and the one shown in Fig. 2. The addition of an avian parasite shows a monophyletic relationship among *P. gallinaceum*, an avian *Plasmodium*, and *P. falciparum*. A phylogenetic comparison using the sequence of the TRAP gene from six different species of *Plasmodium* shows a similar topology and also indicates a monophyletic relationship between *P. gallinaceum* and *P. falciparum* (T. Templeton and D. Kaslow, personal communication). This demonstrates that *P. falciparum* is more closely related to avian parasites than to other human parasites (the *Plasmodium* of chimpanzee, *P. reichenowi*, also falls into this group). It does not suggest that *P. falciparum* arose directly from *P. gallinaceum* but rather that they have a common ancestry. There may presently be a more closely related avian *Plasmodium* species in existence or, alternatively, the similarity may represent a conservation of ancestral features in avian and *P. falciparum* parasites. The possibility of the lateral transfer of parasites between hosts has often been suggested to explain the similarity of parasites within different hosts. Although this probably happens, the direction of such transfers can only be inferred. For example, although it is commonly believed that monkey malarias in the New World arose by transfer of parasites from a human reservoir during colonization, Escalante and Ayala (22) have suggested that colonizers were actually the victims rather than the source. They further suggest that the human parasites *P. malariae* and *P. vivax* originated by way of lateral transfer to humans from New World monkeys at the time of European colonization and that, from these origins, the malarias were subsequently disseminated throughout Europe, Africa, and Asia. Although studies on the distribution of heritable blood types that convey a resistance to *P. vivax* [i.e., Duffy negative blood groups (38)] would tend to lead one to reject this theory, the molecular analysis of *Plasmodium* genes does not support or deny their suggestions.

Molecular data for a common evolutionary ancestry for *P. falciparum* and avian *Plasmodium* species is supported by structural data on the parasites (30). All mammalian *Plasmodium* species have round gametocytes, except *P. falciparum*/*P. reichenowi* that have falciform gametocyte stages. Some of the avian *Plasmodium* species also have elongated falciform gametocytes that fit within elongated avian erythrocytes. The fact that *P. falciparum* gametocytes extend beyond the normal outline of the erythrocyte, distorting its shape, argues for an avian origin for the human parasite. Further evidence for a common ancestry for avian *Plasmodium* and *P. falciparum* was proposed by Sinden *et al.* (30) based on ultrastructural studies on the organelles of gametocytes. We argue from the combined molecular and structural data that the common origin for *P. falciparum* and avian *Plasmodium* may have resulted from transfer of the parasite from evolutionary distant hosts.

We thank E. V. Koonin for suggestions on data analysis, X.-Z. Su for the sequence of the cDNA clones, and D. C. Seeley for technical assistance. J.C.K. was supported by a Fellowship from the National Science Foundation/Sloan Foundation.

1. 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. & Miller, L. H. (1984) *Science* **225,** 593–599.
2. Adams, J. H., Sim, B. K., Dolan, S. A., Fang, X., Kaslow, D. C. & Miller, L. H. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 7085–7089.
3. Cerami, C., Frevert, U., Sinnis, P., Takacs, B., Clavijo, P., Santos, M. J. & Nussenzweig, V. (1992) *Cell* **70,** 1021–1033.
4. Aley, S. B., Bates, M. D., Tam, J. P. & Hollingdale, M. R. (1986) *J. Exp. Med.* **164,** 1915–1922.
5. Chitnis, C. E. & Miller, L. H. (1994) *J. Exp. Med.* **180,** 497–506.
6. Sim, B. K., Chitnis, C. E., Wasniowska, K., Hadley, T. J. & Miller, L. H. (1994) *Science* **264,** 1941–1944.
7. Garnham, P. C. C. (1966) in *Malaria Parasites and Other Haemosporidia* (Blackwell, Oxford), pp. 60–84.
8. Osaki, L. S., Gwadz, R. W. & Godson, G. N. (1984) *J. Parasitol.* **70,** 8321–8323.
9. Dame, J. B. & McCutchan, T. F. (1987) *Exp. Parasitol.* **64,** 264–266.
10. Krettli, A. U., Rocha, E. M. M., Lopes, J. D., Carneiro, C. R. W., Kamboj, K. K., Cochrane, A. H. & Nussenzweig, R. S. (1988) *Parasite Immunol.* **10,** 523–533.
11. Maniatis, T., Fritsch, E. F. & Sambroook, S. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
12. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1992) *Short Protocols in Molecular Biology* (Wiley, New York).
13. McCutchan, T. F., Hansen, J. L., Dame, J. B. & Mullins, J. A. (1984) *Science* **225,** 625–628.
14. Higgins, D. G. & Sharp, P. M. (1988) *Gene* **73,** 237–244.
15. von Heinje, G. (1986) *Nucleic Acids Res.* **14,** 4683–4690.
16. Ramirez, A. D., Rocha, E. M. & Krettli, A. U. (1995) *J. Eukaryotic Microbiol.* **42,** 705–708.
17. Galinski, M. R., Arnot, D. E., Cochrane, A. H., Barnwell, J. W., Nussenzweig, R. S. & Enea, V. (1987) *Cell* **48,** 311–319.
18. Sinnis, P., Clavijo, P., Fenyö, D., Chait, B. T., Cerami, C. & Nussenzweig, V. (1994) *J. Exp. Med.* **180,** 197–306.
19. Goundis, D. & Reid, K. M. B.(1988) *Nature (London)* **335,** 82–84.
20. Robson, J. H., Hall, J. R. S., Jennings, M. W., Harris, T. J. R., Marsh, K., Newbold, C. I., Tate, V. E. & Weatherall, D. J. (1988) *Nature (London)* **335,** 79–82.
21. Rich, K. A., George, F. W., Law, J. L. & Martin, W. J. (1990) *Science* **249,** 1574–1577.
22. Escalante, A. A., Barrio, E. & Ayala, F. J. (1995) *Mol. Biol. Evol.* **12,** 616–626.
23. Waters, A. P., Higgins, D. G. & McCutchan, T. F. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 3140–3144.
24. Arnot, D. E., Barnwell, J. W. & Stewart, M. J. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 8102–8106.
25. McCutchan, T. F., Lal, A. A., do Rosario, V. & Waters, A. P. (1992) *Mol. Biochem. Parasitol.* **50,** 37–46.
26. Ramirez, A. D., Rocha, E. M., Krettli, A. U. (1991) *J. Protozool.* **38,** 40–44.
27. Grassi, B. & Feletti, R. (1890) *Arch. Ital. Biol. (Turin)* **13,** 297–300.
28. Hartman, E. (1927) *Arch. Protistenkd.* **60,** 1–7.
29. Sergent, E., Sergent, E. & Catanei, A. (1929) *Arch. Inst. Pasteur Alger.* **7,** 223–238.
30. Sinden, R. E., Canning, E. U., Bray, R. S. & Smalley, M. E. (1978) *Proc. R. Soc. London* **201,** 375–399.
31. McCutchan, T. F., Dame, J. B., Miller, L. H. & Barnwell, J. (1984) *Science* **225,** 808–811.
32. Brooks, D. R. & McLennan, D. A. (1992) *J. Parasitol.* **78,** 564–566.
33. Ayala, F. J. & Fitch, W. M. (1992) *Parasitol. Today* **8,** 74–75.
34. Siddall, M. E. & Barta, J. R. (1992) *J. Parasitol.* **78,** 567–568.
35. Corredor, V. & Enea, V. (1993) *Mol. Biol. Evol.* **10,** 924–926.
36. Escalante, A. A. & Ayala, F. J. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 11373–11377.
37. McCutchan, T. F., Li, J., McConkey, G. A., Rogers, M. J. & Waters, A. P. (1995) *Parasitol. Today* **11,** 134–138.
38. Miller, L. H., Masson, S. J., Dvorak, J. A., McGinnis, M. H. & Rothman, I. K. (1976) *Science* **189,** 561–563.