Proteolytic Processing and Primary Structure of *Plasmodium falciparum* Apical Membrane Antigen-1*

Received for publication, April 6, 2001, and in revised form, June 1, 2001
Published, JBC Papers in Press, June 8, 2001, DOI 10.1074/jbc.M103076200

Steven A. Howell‡, Chrislaine Withers-Martinez‡§§, Clemens H. M. Kocken, Alain W. Thomas¶, and Michael J. Blackman**

From the §Division of Protein Structure and the ¶Division of Parasitology, National Institute for Medical Research, Mill Hill, London NW7 1AA, United Kingdom and the ‡Biomedical Primate Research Centre, Lange Kleiweg 157, 2280 GJ Rijswijk, The Netherlands

*This work was supported by the Medical Research Council and by European Commission Key Action 2 Quality of Life Program Contract CT-99-01293 and INCO-DEV CT98-0396 (to A. W. T.).

‡ On sabbatical scientific leave from the CNRS-UPR9039 (AFMB), IBSM 31 Chemin Joseph-Aiguier, 13402 Marseilles cedex 20, France. Supported by the Medical Research Council.

** To whom correspondence should be addressed. Tel.: 44 208 959 3666 (ext. 2127); Fax: 44 208 913 8593; E-mail: mblackm@nimr.mrc.ac.uk.

Red cell invasion is critical to survival of the blood stage malaria parasite and is an important potential target for vaccines or drugs designed to block multiplication of the parasite. Despite this, many aspects of invasion remain obscure. In particular, the molecular interactions that mediate and control the sequential processes of host cell binding, merozoite reorientation and junction formation, parasitophorous vacuole formation, and host cell entry are only just beginning to be dissected (1, 2). A number of merozoite proteins known to play a critical role in invasion or both.

**To begin to elucidate the functional significance of the processing, we have performed a detailed structural study of affinity-purified, parasite-derived PfAMA-1. We have first defined precisely the site at which the protein is processed to produce its 66-kDa form. From this information, it is now possible to accurately define the boundary between the proteolytic event that involves the loss of an N-terminal segment and conversion to a 66-kDa form. Both remain membrane-bound, but remarkably, the 66-kDa form is able to redistribute onto the surface of the merozoite, while the unprocessed 83-kDa form remains apically restricted (4, 19). A similar realocalization of AMA-1 has also been shown to occur in *Plasmodium knowlesi* (3, 8) and *T. gondii* (7, 17). The significance of the proteolytic processing of PfAMA-1 is unknown, but its conserved nature and its apparent association with redistribution of the protein suggest that it may be required for PfAMA-1 to move onto the parasite surface or to mediate its role in invasion or both.

This paper is available on line at http://www.jbc.org

1 The abbreviations used are: AMA-1, apical membrane antigen-1; PfAMA-1, *P. falciparum* AMA-1; PrAMA-1, *P. reichenowi* AMA-1; TgAMA-1, *T. gondii* apical membrane antigen-1; mAb, monoclonal antibody; PNase F, *N*-glycosidase F; PNGase A, peptide-N-glycosidase A; PAGE, polyacrylamide gel electrophoresis; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight; CHAPS, 3-[3-cholamidopropyl]dimethylammonio]-1-propanesulfonic acid.
the protein that are specifically shed following merozoite release and show that this results from further processing events that release most of the PfAMA-1 ectodomain from the merozoite surface. The remainder of the protein is probably carried into the invaded host cell. These observations allow us to localize the epitope recognized by a monoclonal antibody, which prevents erythrocyte invasion to domain I to II. Finally, we have shown that PfAMA-1 is not detectably post-translationally modified by oligosaccharide addition.

**EXPERIMENTAL PROCEDURES**

**Materials and Sequence Analysis—**Trypsin (modified sequencing grade) and Glu-C was from Promega (Madison, WI). Sequencing grade chymotrypsin, recombinant N-glycosidase F (PNGase F), and peptide-N-glycosidase A (PNGase A; from sweet almond) were from Roche Molecular Biochemicals. [35S]Methionine/cysteine (Pro-mix-28) was from Amersham Pharmacia Biotech, as were all chromatography media and low molecular weight marker proteins for SDS-PAGE. 14O water was obtained from Isotec (CR Gas Products Ltd., UK). Prestained high molecular weight marker proteins were from Life Technologies, Inc.

Protease inhibitors p-hydroxymercuribenzoate, phenylmethylsulfonyl fluoride, N-p-tosyl-L-lysine chloromethyl ketone, bestatin, peptatin A, leupeptin, aprotinin, and antipain were all obtained from Sigma. Colloidal Blue protein stain was from NOVEKS (Badalson, CA). Protein sequence alignments and pl and molecular mass predictions were performed using Lasergene DNASTar, version 4.05. Homology searches of the incomplete P. falciparum and other Plasmodium genome data bases were performed with the Basic Local Alignment Search Tool (BLAST) program of the National Center for Biotechnology Information (NCBI), National Institutes of Health, using the CustomBLAST server at the NCBI Web site (www.ncbi.nlm.nih.gov/Malaria/plasmodiumblacus.html).

**Parasites and Antibodies—**In vitro culture and synchronization of erythrocytic forms of P. falciparum (clone 3D7) was as previously described (20, 21). Mature, segmented schizonts were isolated from synchronized (22) labeled with [35S]methionine/cysteine (22). Isolation of young ring stage for invasion experiments. Some were first biosynthetically pulse-radio-labeled by the manufacturers. Briefly, purified antibody was trans-}

N

pled to a 1-ml column of G-Sepharose affinity chromatography. Purified mAb 4G2dc1 was coupled to a 2-ml precolumn of Sepharose 4B. All chromatography was performed at 4 °C and low molecular weight marker proteins for SDS-PAGE. 18O water from Amersham Pharmacia Biotech, as were all chromatography media—glycosidase A (PNGase A; from sweet almond) were from Roche was obtained from Isotec (CK Gas Products Ltd., UK). Prestained high molecular weight marker proteins were from Life Technologies, Inc.

Protease inhibitors p-hydroxymercuribenzoate, phenylmethylsulfonyl fluoride, N-p-tosyl-L-lysine chloromethyl ketone, bestatin, peptatin A, leupeptin, aprotinin, and antipain were all obtained from Sigma. Colloidal Blue protein stain was from NOVEKS (Badalson, CA). Protein sequence alignments and pl and molecular mass predictions were performed using Lasergene DNASTar, version 4.05. Homology searches of the incomplete P. falciparum and other Plasmodium genome data bases were performed with the Basic Local Alignment Search Tool (BLAST) program of the National Center for Biotechnology Information (NCBI), National Institutes of Health, using the CustomBLAST server at the NCBI Web site (www.ncbi.nlm.nih.gov/Malaria/plasmodiumblcus.html).

**Parasites and Antibodies—In vitro culture and synchronization of erythrocytic forms of P. falciparum (clone 3D7) was as previously described (20, 21). Mature, segmented schizonts were isolated from synchronized schizonts and analysis by immunoprecipitation (if radiolabeled) or applied to a mAb 4G2dc1 affinity column. The column was washed with 20 ml Tris-HCl, 400 ml NaCl, pH 7.6, and then bound proteins were eluted with 0.2 m glycine-HCl, 0.15 m NaCl, pH 2.7. No detergent was used during the column purification procedure. Eluate fractions (1 ml) were immediately neutralized by the addition of 1 ml Tris-HCl, pH 9.0, and stored at −70 °C.

Sheared, soluble forms of PfAMA-1 were affinity-purified from parasite culture supernatants. Percoll-enriched mature schizonts (unlabeled or biologically radiolabeled with [35S]methionine/cysteine) were washed, and resedimented at 4°C for 30 min to remove the presence of fresh erythrocytes to allow merozoite release and reinvacion (final hematoctrit ~2%). Culture supernatants were then harvested, clarified by centrifugation, passed through a 0.22-μm filter, and then analyzed by immunoprecipitation (if radiolabeled) or applied to a mAb 4G2dc1 affinity column. The column was washed with 20 ml Tris-HCl, 400 ml NaCl, pH 7.6, and then bound proteins were eluted with 0.2 m glycine-HCl, 0.15 m NaCl, pH 2.7. No detergent was used during the column purification procedure. Eluate fractions (1 ml) were immediately neutralized by the addition of 1 ml Tris-HCl, pH 9.0, and stored at −70 °C.

**Protein Alkylation, SDS-PAGE, and In-gel Proteolytic Digestion—**Affinity-purified proteins were concentrated by centrifugation in 12-kDa molecular mass cut-off Vivaspin Micro ultrafiltration units (Sartorius) in reducing SDS-PAGE buffer (25) and alkylated with iodoacetamide as previously described (25). Proteins were subjected to SDS-PAGE and then either transferred directly to polyvinylidene difluoride membrane for N-terminal amino acid sequencing or stained with Colloidal Blue stain and processed by in-gel proteolytic digestion using a modification of the method described previously (22). Briefly, excised SDS-PAGE protein bands were washed with 60% acetonitrile, 200 ml ammonium bicarbonate, dehydrated with 100% acetonitrile, and then dried and reswollen in a minimal volume of trypsin or chymotrypsin at 2 ng ml−1 in 5 ml ammonium bicarbonate or at 4 μg ml−1 Glu-C in 5 ml sodium phosphate, pH 7.8. In some cases, digestion buffers were prepared using 50% (v/v) 18O water.

After overnight digestion at 32 °C, the supernatant was acidified by the addition of a one-tenth volume of 4% trifluoroacetic acid before analysis by mass spectrometry.

**Glycosidase Treatment—**Proteins to be treated with PNGase F were first denatured by heating at 95 °C for 6 min in 50 μl of 0.4% (w/v) SDS in 16 mM Tris-HCl, pH 7.6, 1% (v/v) β-mercaptoethanol. The sample was diluted with 50 μl of distilled water and then further supplemented with 100 μl of 2× PNGase F digestion buffer (2% (w/v) CHAPS in 40 mM sodium phosphate, pH 6.8 (v/v) EDTA, 2% (v/v) β-mercaptoethanol) and divided into two aliquots. One of these was added 2 units (2 μl) of PNGase F, and then both were incubated overnight at 37 °C prior to analysis by SDS-PAGE. For treatment of peptides with PNGase A, in-gel protease digests were lyophilized and taken up in 40 μl of 5 mM ammonium acetate buffer, pH 5.0. The sample was heated at 95 °C for 1 min to inactivate residual protease activity and then divided into two. One of these was supplemented with 30 micromolar (0.6 μl) of PNGase A, and then both samples were incubated for 24 h at 37 °C before acidification and analysis by mass spectrometry.

**Edman Degradation and Mass Spectrometry—**N-terminal sequence analysis of polyvinylidene difluoride-immobilized proteins was carried out at the Protein and Nucleic Acid Chemistry Facility (Cambridge University) by automated Edman degradation on an Applied Biosystems Procise protein sequencer. Peptide mass fingerprinting (26) was performed using a Reflex-III matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany), equipped with a 337-nm nitrogen laser and a Scout-384 probe, to obtain positive ion mass spectra of digested protein with pulsed ion extraction in reflectron mode. An accelerating voltage of 26 kV was used with detector bias gating set to 2 kV and a mass cut-off of m/z 650. Matrix surfaces were prepared using recrystallized α-cyano-4-hydroxycinnamic acid and nitrocellulose using the fast evaporation method (27). Acidified digestion supernatant (0.4 μl) was deposited on the matrix surface and allowed to dry prior to desalting with water. Trypsin autolysis products were used for internal calibration. Peptide mass fingerprints obtained were searched against the reference protein database presented in the public part of the National Center for Biotechnology Information (NCBI), using the program MASCOT (28). Subsequently, the internal calibration was refined using selected intense PfAMA-1 ions. Matched peptide masses were thus finally obtained with a mean mass accuracy of 7 ppm. In estimation of sequence coverage, any observed masses that matched more than one possible peptide sequence were completely excluded from the analysis.
N-terminal Amino Acid Sequencing of AMA-1 Processing Products Affinity-purified from Schizont Extracts

The N-terminal amino acid sequence of the 83-, 66-, and 46-kDa PfAMA-1-derived proteins (referred to hereafter as PfAMA-1\textsubscript{83}, PfAMA-1\textsubscript{66}, and PfAMA-1\textsubscript{46}, respectively) was determined by Edman degradation. The deduced sequence of the PfAMA-1\textsubscript{83} species corresponds to the full-length PfAMA-1 product derived by removal of the secretory signal peptide from the primary translation product.

The results of the Edman degradation are shown in Table 1. The sequence obtained for PfAMA-1\textsubscript{83} is Gln/Ser-Asn-Tyr-Tyr/Pro, which is consistent with the predicted cleavage site between Gly\textsubscript{24} and Gln\textsubscript{25} (SignalP program). This cleavage site would produce a protein with the N-terminal sequence Gln-Asn-Tyr.

The sequences of PfAMA-1\textsubscript{66} and PfAMA-1\textsubscript{46} were also determined and are shown in Table 1. The sequences are consistent with the predicted cleavage sites between Gly\textsubscript{24} and Gln\textsubscript{25} for PfAMA-1\textsubscript{66} and between Gln\textsubscript{25} and Tyr\textsubscript{26} for PfAMA-1\textsubscript{46}.

The results of the Edman degradation are shown in Figure 1A. The N-terminal sequences of the three major polypeptides, obtained by Edman degradation, are shown alongside in single-letter amino acid code. The extreme left- and right-hand lanes contain molecular mass marker proteins phosphorylase b (94 kDa), bovine serum albumin (67 kDa), ovalbumin (43 kDa), and chymotrypsinogen (30 kDa). Multiple alignment of the N-terminal region of AMA-1 of a number of Plasmodium species and T. gondii is shown (Fig. 1B).
Edman degradation of PfAMA-1<sub>66</sub> and PfAMA-1<sub>83</sub> provided much more definitive data (phenylhydantoin yields were ~10 and 1.25 pmol, respectively). For both proteins, the N-terminal sequence obtained was Ile-Glu-Ile-Val-Glu (Fig. 1A), which corresponds to Ile<sup>97</sup>-Glu<sup>101</sup> of the PfAMA-1 sequence (Fig. 1B). Conversion of the 83-kDa PfAMA-1 precursor to the 66-kDa form therefore involves proteolytic cleavage between Ser<sup>96</sup> and Ile<sup>97</sup> within the motif 92NLFSS | EIIVE<sup>101</sup>. It may be significant that this processing site is not closely flanked by any polymorphic residues in known PfAMA-1 sequences (15, 16). It is also very similar to the corresponding flanking sequence in PrAMA-1 (16), the only difference here being an Ile → Asn substitution at the P1’ position relative to the cleavage site (Fig. 1B). Since cleavage site recognition by proteases is usually determined primarily by residues N-terminal to the scissile bond (i.e. the P4–P1 positions), the absolute conservation of these residues between the two species suggests that processing of both PfAMA-1 and PrAMA-1 probably occurs at the analogous position. Intriguingly, however, there is no obvious homology between this region of PfAMA-1 or PrAMA-1 and AMA-1 sequences from any other Plasmodium species or T. gondii (Fig. 1B).

**Mass Spectrometric Peptide Fingerprinting of the Schizont-derived PfAMA-1 Polypeptides PfAMA-1<sub>66</sub> and PfAMA-1<sub>83</sub>**

MALDI-TOF mass spectrometric analysis of tryptic, chymotryptic, or Glu-C proteolytic digests of all three proteins following reduction and alkylation provided data that supported the above sequencing results as well as additional information on the primary structure of the proteins. In all cases, searching of the obtained peptide masses against the nonredundant NCBI protein data base using MASCOT identified the P. falciparum 3D7 AMA-1 sequence (GenBank<sup>TM</sup> accession number U65407) as the most significant hit (typical MASCOT probability score >300). The MALDI-TOF spectra also reproducibly exhibited a complete absence of any major unidentified peaks, confirming the absence of any other protein contaminants (Fig. 2).

Examination of tryptic digests of PfAMA-1<sub>66</sub> and comparison with the expected masses of tryptic peptides of the 3D7 PfAMA-1 sequence identified peptides that covered most of the entire deduced sequence (Figs. 2 and 3A). Notably absent from the coverage was any sequence derived from the predicted secretory signal peptide, consistent with the proposal (tentatively made above on the basis of the limited N-terminal sequencing data) that PfAMA-1<sub>66</sub> represents the PfAMA-1 primary translation product minus the secretory signal peptide. This suspicion was finally confirmed by the presence of an ion at m/z 2728.229, which corresponds to the predicted N-terminal tryptic peptide 97QNYWEHYPQNSDVYPINEHR<sup>45</sup> with a pyroglutamate modification of the N-terminal glutamine residue (expected m/z 2728.245). This ion was continuously absent from tryptic digests of any of the other PfAMA-1-derived fragments (Fig. 2, and see below). The presence of an N-terminal pyroglutamate modification of the affinity-purified PfAMA-1<sub>66</sub> explains the poor yields obtained during Edman degradation of the protein. The sequence covered by the MALDI-TOF analysis also included an ion at m/z 1636.805, corresponding to the peptide 969ASHHTTPYVLEKPYY<sup>622</sup> (predicted m/z 1636.805), the expected tryptic fragment derived from the extreme C terminus of PfAMA-1. No signal corresponding to a mono- or diphosphorylated derivative of this or any other peptide was detected. Between these two regions, the coverage also included three tryptic peptides (Gly<sup>155</sup>-Lys<sup>177</sup>, Asp<sup>291</sup>-Lys<sup>292</sup>, and Cys<sup>380</sup>-Arg<sup>385</sup>) containing three of the six potential N-glycosylation motifs (Asn-Xaa-Ser/Thr) within the 3D7 PfAMA-1 sequence (Fig. 3A). The three remaining potential N-glycosylation sites reside within two predicted tryptic fragments, Asn<sup>371</sup>-Lys<sup>376</sup> and Cys<sup>403</sup>-Lys<sup>417</sup> (Fig. 3). However, further examination of chymotryptic and Glu-C digests of PfAMA-1<sub>83</sub> identified the chymotryptic peptide 413NVKPTCLINNSSY<sup>425</sup> as well as increasing overall coverage of the mature sequence (Gln<sup>25</sup>-Tyr<sup>265</sup>) to 91%. No peptide encompassing the sixth potential N-glycosylation site at Asn<sup>371</sup> was identified in any of the proteolytic digests.

Analysis by MALDI-TOF of proteolytic digests of PfAMA-1<sub>66</sub> was similarly informative (Figs. 2 and 3B). Prominent in the spectrum of tryptic digests was an ion at m/z 758.450, corresponding to the peptide 97IEIVER<sup>102</sup> (predicted m/z 758.441). The corresponding chymotryptic peptide 97IEIVERSNY<sup>105</sup> was also detected as the most intense ion in chymotryptic digests of PfAMA-1<sub>83</sub>. These peptides were not present in digests of PfAMA-1<sub>66</sub> (Fig. 2). Since neither peptide could result from internal proteolysis (Ile<sup>97</sup> is not preceded by a tryptic or chymotryptic cleavage site), these results are fully consistent with the Edman degradation data in confirming Ile<sup>97</sup> as the N-
terminal residue of PfAMA-1$_{66}$. As in the case of PfAMA-1$_{83}$, coverage of the PfAMA-1$_{66}$ sequence extended right through the putative transmembrane sequence to the extreme C-terminal tryptic peptide of the protein, 609ASHTTPVLMEKPYY622, demonstrating unambiguously that both species possess a completely intact cytoplasmic domain. PfAMA-1$_{66}$ presumably translocates to the merozoite surface in this form.

N-terminal sequencing of PfAMA-1$_{46}$ had shown that it possesses an N-terminal sequence identical to that of PfAMA-1$_{66}$, raising the possibility that it is a processing product of the latter, probably derived by a internal proteolytic truncation toward its C-terminal end. This was supported by MALDI-TOF peptide mapping. Analysis of tryptic digests of PfAMA-1$_{46}$ again identified the N-terminal peptide 97IEIVER102 and, with data from chymotryptic digests, extended coverage only as far as peptide392SHGKGYNWGNYNTETQK408 (not shown).

Identification of Soluble PfAMA-1 Processing Products Shed into Culture Supernatants—The characterization of PfAMA-1$_{46}$ suggested that it represents a product of further processing of PfAMA-1$_{66}$, but it was noted by a internal proteolytic truncation toward its C-terminal end. This was supported by MALDI-TOF peptide mapping. Analysis of tryptic digests of PfAMA-1$_{46}$ again identified the N-terminal peptide 97IEIVER102 and, with data from chymotryptic digests, extended coverage only as far as peptide 992SHGKGYNWGNYNTETQK1025 (not shown).

Identification of Soluble PfAMA-1 Processing Products Shed into Culture Supernatants—The characterization of PfAMA-1$_{46}$ suggested that it represents a product of further processing of PfAMA-1$_{66}$, but it was noted by a internal proteolytic truncation toward its C-terminal end. This was supported by MALDI-TOF peptide mapping. Analysis of tryptic digests of PfAMA-1$_{46}$ again identified the N-terminal peptide 97IEIVER102 and, with data from chymotryptic digests, extended coverage only as far as peptide 992SHGKGYNWGNYNTETQK1025 (not shown).
highly purified schizonts that had been allowed to undergo rupture in the absence of significant quantities of fresh erythrocytes (not shown). The affinity-purified soluble proteins, subsequently referred to as PfAMA-1-44 and PfAMA-1-48, respectively, were analyzed by peptide mass fingerprinting in a similar manner to the schizont-derived products. Mass Spectrometric Peptide Fingerprinting of the Soluble, Shed Polypeptides PfAMA-1-48 and PfAMA-1-44—Mass spectra of tryptic digests of affinity-purified PfAMA-1-48 and PfAMA-1-44 contained a strong signal corresponding to the peptide 97IEIVER102, and in chymotryptic digests the peptide 97IEIVERN102 was detected as the most intense peak. PfAMA-1-48 and PfAMA-1-44 thus both possess the same N terminus as PfAMA-1-66, consistent with them both being derived from processing of the latter. In marked contrast to the situation with PfAMA-1-66, however, the digests contained no peptides deriving from the putative transmembrane or cytoplasmic domains of the sequence. In the case of PfAMA-1-48, coverage extended to residue Lys508 (Fig. 3C), whereas coverage of PfAMA-1-44 extended only as far as residue Tyr425 (Fig. 3D). Interestingly, Tyr425 lies just downstream of the last cysteine of domain II (Cys419), whereas Lys508 is only one residue upstream of Cys509, the most C-terminal cysteine of domain III (18). It was not possible to definitively assign the C terminus of either PfAMA-1-48 or PfAMA-1-44. However, determination of the disulfide bonding pattern of the molecule has shown that all 16 cysteine residues of PfAMA-1 form intramolecular disulfide bonds (18), so it is most likely that PfAMA-1-48 encompasses Cys509 (note that this residue would lie within the predicted tryptic peptide 509SVKPTCLINNSSY425 in the PfAMA-1-44 digest and the tryptic peptide F504KVER512 in the PfAMA-1-48 digest) displayed a clear 18O-containing isotope pattern (not shown). Collectively, these results indicate that in each case the true C terminus of the respective PfAMA-1 fragments lie just downstream of these sequences.

It was noticeable that, when isolated from freshly collected culture supernatants, PfAMA-1-48 was reproducibly slightly more abundant than the PfAMA-1-44 fragment; this was evident both by fluorography and Colloidal Blue staining (Fig. 4, A and B). One explanation of this might be that the smaller fragment is a processing product of the larger fragment and that the proteolysis that produces it occurs subsequent to release of PfAMA-1-48 into culture supernatant. If this were so, it would be predicted that extended incubation of harvested culture supernatants would result in further conversion of PfAMA-1-48 to PfAMA-1-44. However, this was found not to be the case (Fig. 4C). Immunoprecipitation from harvested culture supernatants that had been incubated for an additional period of up to 48 h at 37°C did not result in any discernible change in the relative amounts of the two fragments. Taken together with the peptide mapping data, the soluble nature of these proteins, and their appearance in culture supernatants following schizont rupture, our results suggest that PfAMA-1-48 and PfAMA-1-44 are products of distinct proteolytic processing events in which PfAMA-1-48 is cleaved at either of two possible sites, releasing the bulk of the ectodomain from the parasite surface. The released fragments are then stable to further degradation. No fragment of the size of PfAMA-1-48 was present in culture supernatants; the fact that this species was found only in detergent extracts of schizonts, and only then in very limited amounts, leads us to suspect that it most likely represents a product of limited, artifactual degradation of PfAMA-1-48.

It has been suggested that AMA-1 may function during erythrocyte invasion by binding to receptors on the host cell surface (31). Since PfAMA-1-48 encompasses most of the PfAMA-1 ectodomain, experiments were performed to investigate whether this or PfAMA-1-44 possess erythrocyte binding activity. Culture supernatants containing released, radiolabeled PfAMA-1-48 and PfAMA-1-44 were incubated for 2 h at 37°C with varying amounts of washed human erythrocytes and then analyzed for depletion of the PfAMA-1 fragments. No depletion was observed (not shown). Similarly, subsequent analysis of the erythrocytes by immunoprecipitation with mAb 4G2dc1 showed no evidence for binding of the PfAMA-1 fragments (not shown). It was concluded that, under these conditions at least, neither of the shed PfAMA-1 fragments possess detectable erythrocyte binding capacity.

PfAMA-1 Is Not Extensively N-Glycosylated—MALDI-TOF analysis of proteolytic digests of PfAMA-1-63 and its processed

FIG. 4. Identification of soluble PfAMA-1 fragments which are shed following schizont rupture. A, mAb 4G2dc1 was used to immunoprecipitate proteins from Triton X-100 lysates of biosynthetically radiolabeled schizonts (left-hand lane) or from supernatants of cultures in which labeled schizonts had been allowed to undergo rupture and merozoite release/reinvagination (right-hand lane). Immunoprecipitated proteins were detected by fluorography. B, PfAMA-1-48 and PfAMA-1-44 fragments affinity-purified from culture supernatants on a 4G2dc1 affinity column were subjected to SDS-PAGE on a 10% gel and detected by staining with Colloidal Blue. Positions of the PfAMA-1-derived species are shown. The left-hand lane contains ~1 μg each of molecular weight marker proteins phosphorylase b (94 kDa), bovine serum albumin (67 kDa), ovalbumin (43 kDa), chymotrypsinogen (30 kDa), and soybean trypsin inhibitor (20 kDa) (top to bottom). C, harvested culture supernatants containing PfAMA-1-48 and PfAMA-1-44 fragments shed following rupture of biosynthetically radiolabeled schizonts were incubated further for various times at 37°C prior to analysis by immunoprecipitation. No change in the protein profile was observed over time.
products unambiguously identified peptides encompassing five of the six potential N-glycosylation sites within the deduced sequence. This cannot be taken as definitive evidence for a complete lack of glycosylation at these sites, since N-glycosylation of proteins is often heterogeneous, and low levels of glycopeptides may not be detected in MALDI-TOF as they can often yield relatively weak signals. However, in most cases where N-glycosylation of a protein is heterogeneous, the resulting heterogeneity in molecular mass is evident upon SDS-PAGE analysis. This was not so for any of the authentic PfAMA-1 processed products, which all migrated as distinct single bands on SDS-PAGE. Nonetheless, since no peptide encompassing the sixth potential N-glycosylation site at Asn371 was identified in any of the proteolytic digests, it remained a possibility that this site at least might be modified. To address this, we examined the effects of treating the various PfAMA-1 processing products with PNGase F. Fig. 5 shows that PNGase F treatment of the PfAMA-1 products had no detectable effect on their mobility on SDS-PAGE. The mock-treated and PNGase F-treated affinity-purified PfAMA-144 and PfAMA-148 fragments were then subjected to in-gel digestion with trypsin or chymotrypsin, and the digests were examined by MALDI-TOF. There was no detectable difference between the spectra; of particular significance, the tryptic digests contained no peptide corresponding to NKDASMIK or KDASMIK, the product that would result from PNGase F-mediated cleavage of an N-glycosylated glycan from the tryptic peptide 369NKNASMIK376 or 371NASMIK376 (not shown).

N-Glycosylation of blood stage P. falciparum proteins is thought to be rare (32–35); indeed, no malarial N-linked glycan has been experimentally identified to date. It is conceivable, then, that any N-glycans that do exist may take a form similar to those found in invertebrate and plant glycoproteins, which contain a fucose in an α1,3-linkage to the reducing terminal N-acetylglucosamine. These structures are insensitive to cleavage by PNGase F (36) but can be released by PNGase A, which also requires its substrate to be in the form of small peptides (37). Accordingly, chymotryptic and tryptic digests of affinity-purified PfAMA-148 and PfAMA-144 were treated with PNGase A, and the products were analyzed by MALDI-TOF. Once again, there was no detectable difference between the spectra (not shown). Collectively, our data strongly suggest that

PfAMA-1 is not substantially modified, if it is modified at all, by the addition of conventional N-linked oligosaccharides.

A C-terminal Fragment of PfAMA-1 May Be Carried into the Invaded Red Blood Cell—The identification of large amounts of PfAMA-148 and PfAMA-144 in culture medium indicated that extensive processing and shedding of the PfAMA-1 ectodomain from the merozoite surface must take place prior to erythrocyte invasion. The MALDI-TOF data suggested that the cleavage that releases PfAMA-144 takes place between domain II and domain III and that cleavage to produce PfAMA-148 takes place at a position between the end of domain III and the membrane-spanning sequence. We hypothesized that this would probably result in two different sized C-terminal fragments of PfAMA-1, each containing the transmembrane domain and cytoplasmic sequence, being transported into the newly invaded host cell on the merozoite surface. In support of this idea, Narum and Thomas (4) have previously observed that mAb 28G2dc1, which recognizes an epitope within the extreme C-terminal 11 amino acid residues of the PfAMA-1 cytoplasmic domain, is reactive in immunofluorescence assays with acetone-permeabilized newly invaded ring stages. In an attempt to identify these predicted processing products, the following experiment was performed. Mature schizonts metabolically labeled with [35S]methionine/cysteine were recultured for 7 h in the presence of a 10-fold excess of fresh red cells to allow merozoite release and red cell invasion. The cultures were harvested, and residual schizonts were thoroughly removed from the cultures by multiple rounds of centrifugation over Percoll cushions, as described previously (20). The final preparations, which contained only uninfected red blood cells and young, ring stage parasites, were detergent-solubilized and analyzed by immunoprecipitation with mAb 4G2dc1 and mAb 28G2dc1 in parallel with extracts of the starting schizont preparation. Only traces of the PfAMA-148 and PfAMA-144 polypeptides could be precipitated from the ring extracts with mAb 4G2dc1 (not shown), indicating that they are not carried into the invaded erythrocyte to any substantial degree. Fig. 6 shows that, as expected, mAb 28G2dc1 precipitated both PfAMA-143 and PfAMA-146 from radiolabeled schizont extracts. Traces of these species were also precipitated.
from ring extracts. However, the ring extracts additionally contained two low molecular mass species reactive with mAb 28G2dc1 (arrows in Fig. 6), neither of which was present in the schizont extracts. Extended exposure was required to detect these fragments, suggesting that they are very minor components of the ring extracts. They may be rapidly degraded following invasion. The predicted mass of a PfAMA-1 fragment extending from the end of domain II to the C terminus (Ile246–Tyr622) is about 22,700 Da, so it is possible that one of these species could represent the C-terminal product of PfAMA-1 44 release. Shedding of PfAMA-1 48 would produce a C-terminal fragment extending from the end of domain III to the C-terminal residue; this would be very small (on the order of 9–13 kDa) and would probably not resolve well on conventional SDS-PAGE. It is concluded that C-terminal fragments of PfAMA-1 resulting from shedding of PfAMA-1 44 and PfAMA-1 48 may be carried intact into the host erythrocyte but that the majority of the PfAMA-1 ectodomain is shed prior to or at invasion.

**DISCUSSION**

An ever present difficulty in analyzing proteolytic processing in complex biological systems is that of distinguishing authentic, biologically relevant events from spurious, artifactual proteolysis arising during invasive experimental procedures such as detergent extraction and protein purification. Conversion of PfAMA-1 83 to the PfAMA-1 66 form and selective translocation of the latter to the merozoite surface, has been well documented (4, 19). In the present study, both species were evident in detergent extracts of mature schizonts, as previously observed by several investigators (4, 19, 23). We also reproducibly detected a third, much less abundant species in schizont extracts, here termed PfAMA-1 46. However, this fragment was not observed at all under physiological conditions of schizont rupture and merozoite release in *in vitro* culture; instead, two distinct products were seen, PfAMA-1 44 and PfAMA-1 48, leading us to conclude that PfAMA-1 46 is most likely an artifact of spurious proteolytic degradation during detergent extraction. The remainder of this paper focuses on our characterization of what we consider to be the authentic PfAMA-1 processing products, and their relevance to the function of AMA-1 during erythrocyte invasion.

Conversion of PfAMA-1 83 to PfAMA-1 66 involves an N-terminal truncation resulting from cleavage between Ser216 and Ile226; the two proteins are otherwise structurally indistinguishable. Previously, the first residue of domain I has been defined as Cys149, the first cysteine residue of the mature sequence (16, 18). Our new data now allow us to accurately redefine the beginning of domain I as Ile97. Our results also demonstrate conclusively that the large change in mobility on SDS-PAGE resulting from the processing step, an apparent shift of about 16 kDa, is due only to the loss of 71 amino acid residues (Gln25–Ser216), with a calculated molecular mass of about 7.8 kDa, from the N terminus of PfAMA-1 83. The predicted mass of the entire peptide component of PfAMA-1 83 is 69,185 Da, whereas that of PfAMA-1 66 is 60,481 Da. Thus, while both species exhibit aberrant migration on SDS-PAGE, the PfAMA-1 prosequence clearly has an unusually profound effect on migration of PfAMA-1 83. This may be a function of the acidic nature of the prosequence, which has a predicted pI of 4.49. PfAMA-1 shares a very similar prosequence (Fig. 1B) and undergoes a similar post-translational conversion from an 83-kDa precursor to a 66-kDa form (16), probably as a result of proteolytic cleavage at the corresponding Ser216 residue. All other known *Plasmodium* AMA-1 proteins are synthesized as 66-kDa precursors, and they do not obviously undergo an analogous processing step (see e.g., Refs. 3 and 23). This is despite the fact that (outside of the prosequence region) all of the sequences share significant sequence homology with PfAMA-1 and PrAMA-1 and are of a similar length and amino acid composition. This apparent discrepancy in the post-translational fate of the AMA-1 of different *Plasmodium* species may now be tentatively resolved in the light both of the data presented here and the data of a recent careful characterization of the T. gondii AMA-1 (TgAMA-1) (7). TgAMA-1 is synthesized as a protein of slightly greater than 65 kDa and is rapidly processed to a 65-kDa species by removal of a small fragment from the N terminus (7). Antibodies raised against a peptide corresponding to Ser21–Ser36 of the sequence did not recognize the processed 65-kDa form, suggesting that processing takes place downstream of Ser36 (7). An alignment of AMA-1 amino acid sequences from *T. gondii* and a number of *Plasmodium* species (Fig. 2 and Ref. 7) suggests that Ser36 of the *T. gondii* sequence lies close to the prosequence processing site of PfAMA-1. Moreover, it suggests that the prosequence of the *T. gondii* protein is considerably smaller than that of either PfAMA-1 or PrAMA-1 but is about the same size as that of all other known *Plasmodium* AMA-1 sequences. We propose that all AMA-1 precursors are probably processed by removal of a short N-terminal prosequence and that this is most evident in the case of the *P. falciparum* and *P. reichenowi* AMA-1 simply because of the presence of an insertion that does not exist in the other sequences. Large insertions into otherwise globular proteins are a common feature of *P. falciparum* proteins (38); the significance of this is unclear, but the PfAMA-1 prosequence may represent yet another example of the phenomenon.

MALDI-TOF analysis of proteolytic digests of PfAMA-1 66 showed conclusively that, like the PfAMA-1 83 precursor, it possesses an intact transmembrane and cytoplasmic domain. The C-terminal 55 residues of the *Plasmodium* AMA-1, which form the putative cytoplasmic domain, show a strikingly high degree of conservation between species. Of particular note is the presence of a completely conserved Tyr-Tyr motif at the extreme C terminus of the protein. It has been suggested that this may represent a phosphorylation site (8), but the present study argues against this. Tryptic digests of both PfAMA-1 66 and PfAMA-1 83 contained a strong signal corresponding to the mass of the peptide ASHTTPVLMEKPPYY, the predicted tryptic fragment derived from the C terminus of PfAMA-1, but no signal corresponding to a mono- or diphosphorylated derivative of this peptide. Our data are therefore consistent with those of Narum and Thomas (4), who failed to detect any reactivity between PfAMA-1 and an anti-phosphotyrosine antibody. If, for example, phosphorylation of PfAMA-1 is associated with processing to the 66-kDa form, it might be expected that at least in the latter case it should have been possible to detect phosphorylated forms of the peptide.

Following its translocation to the merozoite surface, PfAMA-1 66 is further processed to produce two fragments, PfAMA-1 44 and PfAMA-1 48, which are released in a fully soluble form into culture supernatants. These species result from processing of PfAMA-1 66 at either of two distinct sites, one probably lying between domains II and III and the other at a membrane-proximal position at the C-terminal end of domain III (Fig. 7). From the relative yields of the two proteins, cleavage at the membrane-proximal site appears to be slightly favored over cleavage at the alternative site. For this reason this is unclear. It was not possible in this study to definitively identify the C-terminal residues of PfAMA-1 44 and PfAMA-1 48. The mAb used here to identify PfAMA-1 44 and PfAMA-1 48, mAb 4G2dc1, also recognizes the *P. reichenowi* AMA-1 and inhibits erythrocyte invasion by merozoites of both species *in vitro* (23, 16). The epitope recognized by mAb 4G2dc1 may therefore delineate a functionally important region of the molecule.
low the epitope to be mapped to a position between Ile^{97} and Tyr^{425} within domains I to II (the region encompassed by PfAMA-1_{44}). It is informative to compare our findings with those of Deans et al. (3), who studied the post-translational fate of the P. knowlesi AMA-1. These authors found that following schizont rupture the molecule is proteolytically processed to form two soluble fragments of 42 and 44 kDa, which are shed from the merozoite surface. Similar to the present study, the processed products were identified using a mAb (mAb R3/1C2) that had previously been shown to inhibit erythrocyte invasion by P. knowlesi merozoites (39, 40) and that showed no reactivity with newly invaded ring stage parasites (3, 8). It is likely that the 42- and 44-kDa fragments represent the P. knowlesi homologues of PfAMA-1_{44} and PfAMA-1_{48}, and this provides further evidence that antibody binding to domains I to II of AMA-1 can interfere with its function. Additional parallels with the present data are also evident in the T. gondii system. In their study of TgAMA-1, Hehl et al. (7) found that the protein was shed from the surface of free tachyzoites in a soluble, 54-kDa form that was not recognized by an antibody specific for a region of the putative cytoplasmic domain. The 54-kDa form was not found associated with the parasite surface, and the authors concluded that it resulted from cleavage of the ectodomain at a membrane-proximal site (7). In an independent study Donahue et al. (17) reached a similar conclusion, showing that processing of TgAMA-1 resulted in the shedding of a 53-kDa fragment, while a 12-kDa fragment, probably corresponding to the transmembrane and cytoplasmic domain, remained associated with the parasite. The pattern that emerges shows that translocation of AMA-1 to the surface of the parasite, and its subsequent shedding as a result of controlled, specific proteolysis is a conserved process that transcends species barriers. It is likely to be intimately associated with the function of AMA-1. Our data also suggest that the protease mediating processing of the surface form of PfAMA-1 is parasite-derived, since correct processing can apparently take place in the absence of host cells available for invasion. Identification of this enzyme(s) will be a major focus of future work.

N-Glycosylation of *P. falciparum* blood stage proteins is rare, particularly in the latter stages of schizont maturation during which AMA-1 is synthesized (4, 33). It is important to determine experimentally whether PfAMA-1 is glycosylated, since many interactions between eukaryotic cell surface proteins involve carbohydrate side chains. Any such modifications that do exist could therefore be functionally important. In addition, PfAMA-1 is currently of great interest as a potential component of developmental antimalarial vaccines (1, 2, 41). If vaccine development or functional studies of PfAMA-1 are to be performed using recombinant proteins expressed in heterologous systems that do not incorporate N-linked glycans (such as *Escherichia coli*), or that may introduce inappropriate glycosylation at positions that are not modified in the parasite (e.g., yeast or baculovirus; Ref. 25), it is important to know how the recombinant products may differ antigenically or structurally from the authentic parasite protein. The deduced PfAMA-1 primary sequence contains a total of six potential N-glycosylation sites, all of which are situated within domains I–III of the molecule. Multiple alignments of *Plasmodium* AMA-1 sequences (e.g., Ref. 7) demonstrate that none of these potential N-glycosylation motifs are completely conserved across *Plasmodium* species, an observation that alone may be taken to suggest that N-glycosylation is not critical for AMA-1 function. Here we have identified unambiguously peptides comprising five of these motifs. Furthermore, we were unable to demonstrate any susceptibility to PNGase F or PNGase A, strongly suggesting that PfAMA-1 is not modified by the addition of sizable N-linked oligosaccharides. However, in MALDI-TOF analysis, no peptide comprising the sixth potential N-glycosylation site at Asn^{371} could be identified, so it remains possible that this or an adjacent residue within the sequence ^NASMI^ may be modified in some way. This is unlikely to be O-glycosylation, since submission of the 3D7 PfAMA-1 sequence to the prediction server of the Center for Biological Sequence Analysis (www.cbs.dtu.dk/services/) did not identify Ser^{73} as a potential O-glycosylation site. Nonetheless, it may be significant that an examination of all available PfAMA-1 sequences deposited in GenBank™ (a total of 35, derived from distinct parasite isolates) shows that ^NASMI^ and the surrounding sequence is completely conserved. Further work is necessary to investigate whether this region of the protein displays a subtle post-translational modification.

Of what significance might the complex proteolytic processing of PfAMA-1 be for its physiological function? Interest in AMA-1 was first sparked by the demonstration that treatment of isolated *P. knowlesi* merozoites with anti-AMA-1 mAbs or Fab fragments thereof rendered them noninvasive (39, 40). AMA-1 is therefore thought to play a role in invasion, and it presumably requires exposure on the merozoite surface in order to perform this role. The first processing step following transport to the rhoptries, loss of the prosequence, may be a prerequisite or trigger for further mobilization of the protein onto the merozoite surface. Once there, the modified molecule...
could perhaps act in an adhesive role, mediating binding to the host cell surface, as has been suggested by Kappe et al. (31). These authors have identified a family of \textit{Plasmodium} integral membrane proteins called MAEBL, which contain tandem cysteine-rich domains with significant homology to AMA-1 domains I and II. Expression of these MAEBL-derived structures in a membrane-bound form on the surface of COS-7 cells demonstrated that they possessed erythrocyte-binding activity. This finding is now of particular interest given our demonstration that the epitope recognized by the inhibition-inhibitory mAb 4G2dc1 lies within domains I and II of PIAMA-1, the region that shows most similarity to MAEBL. The relative importance of domains I and II may also be reflected in the high degree of similarity between the \textit{Plasmodium} and \textit{T. gondii} AMA-1 sequences within this part of the molecule; domain III is much less conserved (7). There are other intriguing parallels between AMA-1 and adhesins known to be involved in apicomplexan gliding motility and invasion, such as the \textit{T. gondii} tachyzoite protein MIC2 and the \textit{Plasmodium} sporozoite protein TRAP. Both of these molecules are released in a membrane-bound form from apical secretory granules (micronemes) to traverse the surface of the zoite in an anterior-to-posterior direction and then are shed as a result of proteolytic activity, which releases the bulk of the protein from its membrane-spanning domain (42, 43). Our data suggest that PIAMA-1 is subject to a similar fate. The mechanism which drives translocation of PIAMA-1 onto and across the merozoite surface has not been investigated, but the molecule could be linked, via its highly conserved cytoplasmic domain, to a subpellicular actomyosin motor in a similar manner to that proposed for MIC2 and TRAP (43). The present study found no evidence that the shed forms of PIAMA-1 (PIAMA-1a and PIAMA-1c) can bind to red blood cells. However, there is accumulating evidence that adhesins involved in host cell invasion by apicomplexan parasites may act in a cooperative manner, perhaps clustering at the moving junction in the form of homo- or heteromers to allow multivalent, high avidity interactions with host cell receptors. If intrinsic binding affinity is low, it may be difficult to visualize in solution. Alternatively, as in the case of MIC2 (44), proteolytic cleavage of the PIAMA-1 ectodomain may have the effect of reducing the affinity of the interaction between the molecule and its receptor. As proposed by Carruthers et al. (42, 44), this may be necessary for successful completion of invasion, providing a mechanism for eventual disengagement of the ligand-receptor complex as it reaches the posterior pole of the invading merozoite. This hypothesis is now testable; we have recently successfully expressed the entire PIAMA-1 ectodomain at high levels in a soluble, correctly folded recombinant form, and this will provide an invaluable reagent for future studies on the structure and function of the protein.

Specific proteolytic processing and targeted shedding of apicomplexan zoite surface proteins during host cell invasion is now becoming a recurring theme (45). The proteases that mediate these critical events at the surface of the malaria merozoite may provide exciting targets for new generations of protease inhibitor-based drugs, urgently needed to combat the continuing spread of resistance of the parasite to available antimalarials.

Acknowledgements—We are grateful to Fiona Hackett and Peter Fletcher for excellent technical assistance. We also thank Dr. Tony Holder and Professor Guy Dodson for invaluable help, advice, and encouragement.

\*\*\*\*\*\*

2 C. Kocken, C. Withers-Martinez, F. Hackett, M. Dubbeld, A. van der Wel, M. Blackman, and A. Thomas, manuscript in preparation.