The Disulfide Bond Structure of Plasmodium Apical Membrane Antigen-1*

(Received for publication, March 15, 1996, and in revised form, May 29, 1996)

Anthony N. Hodder‡‡‡, Pauline E. C. Crewther†, Mary L. S. M. Matthew‡, Gavin E. Reid‡, Robert L. Moritz‡, Richard J. Simpson‡, and Robin F. Anders‡‡

From the ‡‡ Cooperative Research Centre for Vaccine Technology and the Department of Microbiology, University of Melbourne, Melbourne, Australia 3052, the †† Cooperative Research Centre for Vaccine Technology and The Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia 3050, and the ‡‡ Joint Protein Structure Laboratory, Ludwig Institute for Cancer Research (Melbourne Branch) and The Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia 3050

Apical membrane antigen-1 (AMA-1) of Plasmodium falciparum is one of the leading asexual blood stage antigens being considered for inclusion in a malaria vaccine. The ability of this molecule to induce a protective immune response has been shown to be dependent upon a conformation stabilized by disulfide bonds. In this study we have utilized the reversed-phase high performance liquid chromatography of dithiothreitol-reduced and nonreduced tryptic digests of Plasmodium chabaudi AMA-1 secreted from baculovirus-infected insect cells, in conjunction with N-terminal sequencing and electrospray-ionization mass spectrometry, to identify and assign disulfide-linked peptides. All 16 cysteine residues that are conserved in all known sequences of AMA-1 are incorporated into intramolecular disulfide bonds. Six of the eight bonds have been assigned unequivocally, whereas the two unassigned disulfide bonds connect two Cys-Xaa-Cys sequences separated by 14 residues. The eight disulfide bonds fall into three nonoverlapping groups that define three possible subdomains within the AMA-1 ectodomain. Although the pattern of disulfide bonds within subdomain III has not been fully elucidated, one of two possible linkage patterns closely resembles the cystine knot motif found in growth factors. Sites of amino acid substitutions in AMA-1 that are well separated in the primary sequence are clustered by the disulfide bonds in subdomains II and III. These findings are consistent with the conclusion that these amino acid substitutions are defining conformational disulfide bond-dependent epitopes that are recognized by protective immune responses.

The development of a malaria vaccine is an urgent priority particularly because drug-resistant strains of Plasmodium are becoming more prevalent. Strategies for developing such vaccines are being targeted toward various stages of the Plasmodium lifecycle, particularly the asexual merozoite stage, which invades the host erythrocyte. The intraerythrocytic parasite undergoes asexual reproduction, producing a new generation of merozoites, which become directly accessible to immune attack at the time of schizont rupture. As a consequence, there is a great deal of interest in the identification and characterization of merozoite surface antigens as potential vaccine candidates (1). One such protein is the apical membrane antigen-1 (AMA-1) (2–8), a minor surface antigen that is synthesized in mature, segmenting schizonts during the final 4 h of erythrocytic development (9). Initially, AMA-1 appears to be located in the electron-dense neck of the rhotries, but after schizont rupture a processed form of the protein spreads circumferentially around the surface of mature merozoites (9).

Although the biological function of AMA-1 is unknown, its location and stage specificity suggest that it may be involved in the process of erythrocyte invasion (9, 10). A monoclonal antibody raised against native PK66, the Plasmodium knowlesi homologue of AMA-1, and Fab fragments of this antibody prevented P. knowlesi merozoites from invading rhesus erythrocytes in vitro (11). Furthermore, passive transfer of AMA-1-specific polyclonal antibodies into Plasmodium chabaudi-infected mice prevented lethal parasitemias. Protection against parasite challenge can also be induced by active immunization with AMA-1. For example, immunization of rhesus monkeys with the native PK66, affinity purified from infected rhesus erythrocytes, partially protected against blood stage challenge (10). Recombinant AMA-1 expressed using baculovirus-infected insect cells has protected immunized monkeys and mice against simian and rodent parasites, respectively (12). More recently, mice were protected against infection with P. chabaudi by immunization with the presumed ectodomain of P. chabaudi AMA-1 expressed in Escherichia coli. This fragment of AMA-1 (AMA-1B) lacked the putative transmembrane and C-terminal cytoplasmic domains of the full-length polypeptide (residues 481–503 and 503–558, respectively). Protection was induced with antigen that had been refolded in vitro with the formation of intramolecular disulfide bonds but not by antigen that had been reduced and alkylated. Full sequences have been reported for the AMA-1 genes of 11 isolates of Plasmodium falciparum (3, 4, 13), two strains of P. knowlesi (14), and strains of Plasmodium vivax (7), Plasmodium fragilis (5), Plasmodium cynomolgi (8), and P. chabaudi (6). Unlike many other asexual stage antigens, AMA-1 lacks immunodominant blocks of repetitive sequence and has the

* This work was supported by the Cooperative Research Centre for Vaccine Technology, the United Nations Development Program/World Bank/World Health Organization Special Programme for Research and Training in Tropical Diseases and National Health and Medical Research Council (Australia). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Cooperative Research Centre for Vaccine Technology and The Walter and Eliza Hall Inst. of Medical Research, Post Office Royal Melbourne Hospital, Victoria 3050, Australia. Tel.: 61-3-9345-2555; Fax: 61-3-9347-0852.

3 The abbreviations used are: AMA-1, apical membrane antigen-1; AMA-1B, apical membrane antigen ectodomain; DTT, dithiothreitol; PAGE, polyacrylamide gel electrophoresis; HPLC, high performance liquid chromatography; RP, reversed-phase.

§ R. F. Anderson, P. E. Crewther, S. Edwards, M. Hayden, M. L. S. M. Matthew, B. Pollack, and D. Pye, submitted for publication.
Disulfide Bonds in AMA-1

structural characteristics of a Type I integral membrane protein. The ectodomain of AMA-1B contains 16 cysteine residues, which are conserved within all known sequences of AMA-1. Because a conformation stabilized by disulfide bonds is necessary for AMA-1 to induce a protective immune response, we have sought to determine the disulfide bond topology in P. chabaudi AMA-1 secreted from baculovirus-infected insect cells. For this we used a combination of reversed-phase HPLC peptide mapping under reducing and nonreducing conditions and subsequent N-terminal sequence and mass analysis of the disulfide-linked peptides.

MATERIALS AND METHODS

Expression of P. chabaudi AMA-1 Using Baculovirus—The methods used for cloning AMA-1 sequences in baculovirus and expression in S9 insect cells will be described in detail elsewhere. Briefly, DNA encoding AMA-1B, comprising amino acid residues 1–499 of the 558-residue full-length polypeptide (GenBank accession number U49743), was amplified by polymerase chain reaction from P. chabaudi adami S9 strain DNA. PCR products were cloned into the baculovirus expression vector ArCP23. Protein DNA produced in E. coli mixed with DNA from wild-type baculovirus was used to transfect S9 insect cells. Recombinant baculovirus was selected by plaque purification and then amplified in S9 insect cell cultures. Cultures expressing AMA-1B were detected by Western blot analyses using polyclonal antibodies raised to recombinant P. falciparum AMA-1 expressed in S9 cells.

SDS-PAGE and Immunoblot Analysis—Baculovirus-expressed P. chabaudi (strain DS) AMA-1B samples were electrophoresed under reducing conditions on 10% polyacrylamide gels (SDS-PAGE) (15). For immunoblots, proteins were electrothermally transferred from SDS-PAGE gels onto nitrocellulose filters (16). The filters were incubated in 5% (w/v) nonfat milk powder in phosphate-buffered saline and reacted with polyclonal rabbit antibodies raised against refolded P. chabaudi DS AMA-1B expressed in E. coli. The filters were further incubated with 125I-labeled protein A and autoradiographed.

Purification of AMA-1B—Crude AMA-1B supernatant from baculovirus-infected S9 insect cells was dialyzed overnight at 4°C against 0.5 M NaCl/20 mM Tris-HCl buffer, pH 7.4, with continuous stirring and then centrifuged at 14,500 × g for 20 min at 4°C. Supernatant (250 ml) was pumped at 3 ml/min onto 45 ml of concanavalin A-Sepharose (Pharmacia Biotech Inc., Uppsala, Sweden) packed into a Pharmacia C26/60 C-type column using a Bio-Rad (Richmond, CA) automated Econosystem operated at 4°C. The column was then washed with 10 column volumes of 0.5 M NaCl/20 mM Tris-HCl, pH 7.4. Bound proteins were eluted from the column with the same buffer containing 0.5 M NaCl/20 mM Tris-HCl, pH 7.4, 0.05% (v/v) trifluoroacetic acid and 60% (v/v) acetonitrile (Chromatography-grade acetonitrile, Pierce). The solution was mixed with 50 ml of a 1% (w/v) mixture of 1% (w/v) Triton X-100 and Milli Q water. The solution was incubated overnight at 37°C. The reaction was then stopped by the addition of reducing SDS-PAGE sample buffer (15). Human apotransferrin (Sigma) was used as a positive control and was prepared using 10 µl of a 2 mg/ml solution in a similar manner described for AMA-1B. The negative controls, with or without human apotransferrin, were incubated overnight in digest buffer at 37°C without the N-glycanase enzyme. Samples were analyzed using SDS-PAGE and immunoblot analysis as described above.

Tryptic Digestion of AMA-1B—The RP-HPLC peak fraction containing AMA-1B was reduced to half-volume using a Labconco (Kansas City, MO) Centrivap concentrator. The solution was then buffered to pH 6.5 using a 1:10 dilution of 1 M sodium phosphate, pH 6.5. Sequencing grade modified trypsin (Promega, Madison, WI) in 50 mM acetic acid was added at a trypsin/peptide ratio of 1:25 (w/w). The solution was incubated at 37°C for 5 h. A second aliquot of trypsin (1:25 w/w) was then added, and the mixture was incubated for a further 14 h at 37°C. Digestion was stopped by either snap freezing to −70°C or RP-HPLC. RP-HPLC Analysis of AMA-1B Tryptic Fragments—Narrow bore RP-HPLC of DTT-reduced and nonreduced AMA-1B tryptic digests was used to identify the location of disulfide-linked peptides. RP-HPLC was conducted using the Hewlett-Packard 1050 system and buffers described above. An aliquot containing 2.5 µg of tryptic digest was mixed with an equal volume of 6% guanidinium chloride, 0.2% Tris-HCl, pH 8.4, 0.005 M EDTA and then reduced by the addition of DTT (final concentration, 25 mM) for 60 min at 45°C. A tryptic fingerprint for the reduced digest was then obtained by eluting the peptide fragments from a C18 column (2.1-mm inner diameter × 250 mm, Vydac (The Separations Group, Hesperia, CA)) using a 0–100% buffer B linear gradient over 90 min at a flow rate of 0.1 ml/min. The profile obtained was compared with the tryptic fingerprint obtained for a similar amount of nonreduced digest eluted under identical chromatographic conditions. The disappearance of a peak in the nonreduced fingerprint following reduction indicated the presence of a disulfide bond. The remainder of the nonreduced tryptic digest mixture (120 µg) was fractionated under similar conditions using a 0–100% buffer B linear gradient over 90 min and a Vydac C18 (4.6-mm inner diameter × 250 mm) column operated at a flow rate of 1 ml/min. Disulfide-containing peptide fractions were collected and subjected to Edman degradation and electrospray-ionization mass-spectrometric analysis.

Thermolysin Subdigestion of AMA-1B Tryptic Fragment 58—RP-HPLC peak fraction 58 adjusted to 0.01% (v/v) Tween 20 was reduced to half-volume using a Labconco centrivap concentrator. An equal volume of 1% (w/v) unbuffered NH4HCO3, 4 mM CaCl2 was then added to the solution. Thermolysin (Boehringer Mannheim GmbH, FRG) was added to an estimated 1–2 µg of tryptic peptide 58 in 1% (w/v) NH4HCO3 at an enzyme/peptide ratio of 1:10 and incubated for 2 h at 50°C.

RP-HPLC fractionation of the thermolysin digest was performed at 45°C. Buffer A comprised 0.1% (v/v) aqueous trifluoroacetic acid, whereas buffer B comprised aqueous 60% (v/v) acetonitrile containing 0.1% (v/v) trifluoroacetic acid. The thermolysin digest mixture was fractionated on a C8 column (2.1-mm inner diameter × 100 mm). The solution was mixed with an equal volume of a 0–100% buffer B linear gradient over 12 min at a flow rate of 0.5 ml/min (17). Thermolysin peak fractions 9 and 17 were rechromatographed in a second dimension RP-HPLC buffer system to resolve minor peptide contaminants. In this system, buffer A comprised aqueous 154 mM NaCl, and buffer B comprised 60% aqueous acetonitrile containing 154 mM NaCl. Peptides were eluted from a Brownlee RP-300 column using a 0–100% buffer B linear gradient over 60 min at a flow rate of 0.1 ml/min.

Capillary Column RP-HPLC/Electrospray-Ionization Mass Spectrometry—The Edman degradation of purified recombinant AMA-1B revealed a single polypeptide with an N-terminal sequence consistent with that expected for AMA-1B. Enzymatic Deglycosylation of AMA-1B—Baculovirus-expressed AMA-1B was deglycosylated using a recombinant N-glycanase enzyme (EC 3.2.1.23 and 3.2.1.18) from Genzyme (Cambridge, MA). An aliquot (100 µl) of the serum-exchange purified AMA-1B at pH 4.5 was adjusted to a final concentration of 0.5% (v/v) SDS and 50 mM β-mercaptoethanol, and the protein was denatured by boiling for 5 min. An aliquot (20 µl) of this solution was mixed with 5 µl of 7.5% (v/v) Triton X-100 (United Technologies/Packard, Downers Grove, IL), 0.38 units of N-glycanase enzyme (1.5 µl) and made up to 30 µl with Milli Q grade water (3.5 µl). The solution was incubated overnight at 37°C. The reaction was then stopped by the addition of reducing SDS-PAGE sample buffer (15). Human apotransferrin (Sigma) was used as a positive control and was prepared using 10 µl of a 2 mg/ml solution in a similar manner described for AMA-1B. The negative controls, with or without human apotransferrin, were incubated overnight in digest buffer at 37°C without the N-glycanase enzyme. Samples were analyzed using SDS-PAGE and immunoblot analysis as described above.
trometry—Peptides were separated by on-line rapid reversed-phase capillary HPLC coupled to a triple quadrupole mass spectrometer (Finnigan-MAT model TSQ-700, San Jose, CA) equipped with a Finnigan electrospray ionization source configured as described elsewhere (18, 19). The reversed-phase capillary column used in this study (150 mm × 0.2-mm inner diameter, Vydac C18) was fabricated as described elsewhere (20, 21) and developed with a linear 15-min gradient at 1.6 μl/min from 0 to 90% buffer B, where buffer A was 0.1 M aqueous acetic acid, and buffer B was acetonitrile. The electrospray ionization needle was operated at 24.5 kV. The sheath liquid was 70% methanol/0.1M acetic acid delivered at 3 μl/min via a syringe drive (Harvard Apparatus, South Natick, MA). Nitrogen sheath and auxiliary gases obtained from a boiling liquid nitrogen source were supplied at 20 p.s.i. and 15 units (arbitrary value), respectively. The heated capillary was set at 150°C. Mass spectra were collected every 3 s in centroid mode over the m/z range 300-2000. Peptide molecular masses were calculated using Finnigan BIOMASS software. The PEPMAP component of this software was used to generate a theoretical tryptic digest of the AMA-1B amino acid sequence. This program also performs an iterative combination of all tryptic fragments containing Cys residues, enabling prediction of the sequences and molecular masses for the corresponding peptides linked through a single disulfide bond. N-terminal Sequence Analysis—N-terminal amino acid sequencing of peptides was performed by automated Edman degradation using a Hewlett-Packard model G1005A protein sequencer operating with the routine 3.0 sequencer program. A Hewlett-Packard model HP1090A liquid chromatograph was used for phenylthiohydantoin amino acid analysis.

RESULTS

Purification of AMA-1 Ectodomain (AMA-1B)—Recombinant AMA-1B secreted by baculovirus-infected insect cells was purified by affinity chromatography on concanavalin A, followed by anion-exchange and reversed-phase chromatographies. The final RP-HPLC step resolved four peaks (data not shown). The dominant peak was found to correspond to AMA-1B, which was identified by its reactivity on immunoblots with anti-AMA-1B antibodies and by Edman degradation analysis. N-terminal sequence analysis (data not shown) revealed that the signal sequence was cleaved from the mature exported protein between Cys21 and Ser22 of the predicted primary translation product.

The AMA-1B isolated by RP-HPLC was a closely migrating doublet when analyzed by SDS-PAGE (Fig. 1a, lane 1). Prior incubation of AMA-1B with N-glycanase generated a single species of Mr 57,000 (Fig. 1b, lane 3) smaller than either component of the original doublet (Fig. 1b, lane 4). Thus the observed heterogeneity in RP-HPLC-purified AMA-1B was apparently due to differential glycosylation rather than the existence of sequence variants.

RP-HPLC Analysis of AMA-1B Tryptic Digest—Enzymatic cleavage of AMA-1B with trypsin under slightly acidic condi-
tions to reduce the potential for disulfide bond interchange was performed to generate individual disulfide-linked peptides. To identify disulfide-linked peptides, analytical RP-HPLC analyses were performed on identical portions of the tryptic digest with and without DTT reduction. Peptide fractions found to contain disulfide bonded peptides by Edman degradation and electrospray-ionization mass spectrometry are indicated by an asterisk. Peak numbers correspond to those given in Table I. (See “Materials and Methods” for further details).

**Table I**

| Peak Number | Cysteine connectivitiesa | Amino acid sequenceb | Residues | Mass observed Da | Mass calculated Da |
|-------------|--------------------------|----------------------|----------|-----------------|-------------------|
| Tryptic peaks |                          |                      |          |                 |                   |
| 25          | 1–6                      | CPVMGK (30)          | 94–99    | 3704.3          | 1841.10           |
| 56          | 2–3                      | EHPHEALNDMSLCAK (45) | 149–164  | 2556.9          | 2557.04           |
| 42          | 4–5                      | TCYLILY (48)         | 191–196  |                 |                   |
| 46          | 7–10                     | YCSNEENENQFPCTPEK (60)| 207–225  | 2290.3          | 2291.44           |
| 34          | 8–9                      | KPTCLINDK (98)       | 360–368  |                 |                   |
| 58          | 11, 12, 13, 14, 15, 16  | NFVATALLSSLEAQEFPCD (78) | 369–393  | 6611.5          | 6612.52           |
|             |                          | IFISDDKESLKCPEPTQLT (102) | 420–453  |                 |                   |
| Thermolysin peaks |                    |                      |          |                 |                   |
| from 58     |                          |                      |          |                 |                   |
| 17/2        | 11–14                    | LSSLLEAQEFPCDIYK (101) | 376–393  | 2952.5          | 2953.27           |
| 9/3         | 12, 13–15, 16            | LTQSCCNF (85)       | 438–445  |                 |                   |
|             |                          | FVCNCVEK (104)      | 446–453  | 2729.7          | 2730.15           |
|             |                          | ISDDKESLKCPEPTQ     | 422–437  |                 |                   |

* Numbers refer to relative position of cysteines in amino acid sequence.

* One-letter abbreviation used for amino acids; peptide yields (picomoles), calculated from the values for phenylthiohydantoin amino acids in the first cycle of the Edman degradation, are shown in parenthesis. Sample load: 18% of each tryptic peptide fraction and 100% of each thermolytic peptide fraction was taken for analysis.

* X indicates a glycosylated Asn residue.

* This is a partial sequence and consequently does not correspond to the observed molecular ion mass. The calculated mass was determined by extending this sequence C-terminally to the next tryptic cleavage site (i.e. Lys393 and Lys453).

* Multiple ion series consistent with the presence of carbohydrate.
The observed mass of 2290.3 Da for the peptide in peak fraction 42 corresponded to the single tryptic peptide (Tyr207–Lys225) with an internal disulfide bond between Cys206 and Cys220. The observed mass of the peptide in peak fraction 46 (2708.6 Da) was consistent with tryptic peptides Phe257–Lys270 and Lys360–Lys368 disulfide-linked through Cys360 and Cys368. Edman degradation confirmed the predicted sequences for the peptides involved in these three disulfide bonds.

The remaining disulphide-linked peptides that were not predicted by PEPMAF software because they were either glycosylated (peak 25) or aberrant cleavage products (peak 56 and 58) were identified using a combination of mass analysis and Edman degradation.

A broad peak eluting at 69 min in the analytical RP-HPLC was lost with DTT reduction (Fig. 2). There were two well resolved peaks (56 and 58) in the corresponding position on the preparative chromatogram (Fig. 3). Edman degradation of the tryptic fraction 56 established that two peptides comprising residues Glu149–Lys164 and Thr191–Tyr196 were linked by a disulfide bond involving Cys162 and Cys192. The C-terminal Tyr on peptide Thr191–Tyr196 indicated a chymotrypsin-like cleavage that was responsible for the failure of the PEPMAF program to predict this pair of disulfide bonded peptides. The calculated mass of these two disulfide bonded peptides (2557.04 Da) was in close agreement with the observed molecular ion mass (2556.9 Da).

Edman degradation of fraction 58 revealed two peptide sequences commencing at Asn369 and Ile420, respectively (see Table I). The sequence of the peptide commencing with Ile420 contained two tryptic sites resistant to cleavage (Lys426 and Lys430), and the observed molecular ion mass of 6611.5 Da was consistent with a fraction containing peptides with trypsin-resistant sites. The calculated mass (6612.52 Da) of two peptides Asn369–Lys383, Ile420–Lys430 being the first tryptic potential site C-terminal to Asn369 and Ile420–Lys430, Lys368 being the first potential tryptic site C-terminal to the trypsin-resistant sites Lys426 and Lys430) linked by three disulfide bonds was in close agreement with the observed molecular ion mass (6611.5 Da). Because it was not possible to determine the connectivities of the cysteine residues in the two peptides in fraction 58, this fraction was digested further using thermolysin (see below).

Of the cysteines not assigned to the disulfide bonds discussed above, one (Cys307) was within a predicted tryptic peptide (Glu241–Lys256), which also contained a tryptic fragment (Trp245). A candidate peak fraction (25) containing this peptide was identified by examining the chromatogram monitored at 280 nm (not shown), and Edman degradation established that the tryptic-containing peptide (Glu241–Lys256) was linked to peptide Cys48–Lys99 by a disulfide bond between Cys48 and Cys247.

The observed mass for fraction 25 was more than twice that of the calculated mass of the two disulfide bonded peptides in this fraction (Table I). An inspection of the AMA-1B sequence revealed that Asn249, a potential site for N-glycosylation (Asn-Xaa-Ser), is immediately followed by a site for trypsin cleavage (Lys250). A blank cycle 9 of the Edman degradation of 25 was consistent with glycosylation of Asn249. Mass analysis of peak fraction 25 revealed ions of double and triple charge with a calculated mass of 3704.3 Da. The observed ion series showed a sequential loss of 162 Da for the nine most distal carbohydrate residues followed by two further consecutive losses of 203 Da to give the expected molecular ion mass of 1841.10 Da for the nonglycosylated disulfide-linked peptide. The loss of 162 and 203 Da is consistent with the loss of hexose-like (Fru, Gal, Gla, and Man) and N-acetylgalactosamine-like (GalNAc or GlcNAC) residues, respectively (22).
Gln\textsuperscript{437} and Phe\textsuperscript{446}–Lys\textsuperscript{453} being linked through two disulfide bonds. The Edman degradation data for peptide Phe\textsuperscript{446}–Lys\textsuperscript{453} confirmed the postulated C terminus for peptide Ile\textsuperscript{420}–Lys\textsuperscript{453} discussed above.

Although unequivocal disulfide bond assignments could not be made for Cys residues 431, 433, 448, and 450, thermolysin cleavage data revealed that the Cys\textsuperscript{431}–Cys\textsuperscript{433} and Cys\textsuperscript{448}–Cys\textsuperscript{450} disulfide bond linkages could be excluded. The two unassigned disulfide bonds therefore must connect the two Cys-Xaa-Cys sequences as Cys\textsuperscript{431}–Cys\textsuperscript{448} and Cys\textsuperscript{433}–Cys\textsuperscript{450} or Cys\textsuperscript{431}–Cys\textsuperscript{450} and Cys\textsuperscript{433}–Cys\textsuperscript{448}.

**DISCUSSION**

The disulfide bond-dependent conformation of AMA-1 is essential for inducing a protective immune response against malaria in animal models.\textsuperscript{7} To better understand the specificity of the protective immune response induced by immunization with AMA-1, we have sought to determine the pattern of disulfide bonds in this important vaccine candidate. The parasite antigen is present in low abundance and is only expressed over a narrow time window during the cycle of asexual development in the host erythrocyte, and for this reason we have used recombinant AMA-1 produced using the baculovirus expression system for the disulfide bond assignments. The recombinant antigen used lacked the cytoplasmic and putative transmembrane domains but retained all of the putative ectodomain sequence including all 16 conserved cysteine residues. The cloned DNA construct included the sequence encoding the AMA-1 signal for secretion and the antigen with the signal peptide cleaved was recovered from the medium in which the transfected Sf9 insect cells were cultured. This form of AMA-1 was highly antigenic when reacted with the serum of hyperimmune mice. For this reason, and because the AMA-1 ectodomain was efficiently transported through the secretory pathway of the insect cell and was recovered from the culture supernatant in a glycosylated and stable form, we assumed that the protein was folded with the correct disulfide bonding pattern.

The N-terminal sequencing and mass spectrometric analyses

---

**FIG. 5.** Schematic diagram showing the disulfide bonds and subdomains in *P. chabaudi adami* DS AMA-1. The vertical lines below the schematic indicate the location of amino acid substitutions that occur between *P. chabaudi adami* DS and 556KA AMA-1 sequences.

**FIG. 6.** *a*, a schematic diagram showing the relationship between the sites of amino acid substitution occurring among *P. chabaudi adami* strains DS and 556KA and *P. chabaudi chabaudi* strain CB and the disulfide bonds within AMA-1. Sequence data were obtained from footnote.\textsuperscript{9} Note that *P. chabaudi chabaudi* CB AMA-1 has only been partially sequenced (residues 51–377). The sequence shown is that for *P. chabaudi* DS AMA-1 ectodomain (residues 86–479). *b*, a schematic diagram showing the relationship between the sites of amino acid substitution occurring among eleven *P. falciparum* isolates obtained from Ref. 4. The sequence shown is that of *P. falciparum* 3D7 AMA-1 ectodomain (residues 141–538). Black circles represent disulfide-linked Cys residues, and shaded circles represent the location of amino acid substitutions within the primary sequence. The disulfide bond arrangement between the two Cys-Xaa-Cys motifs in domain III has not been resolved; however, the connectivities portrayed are considered the most likely.
of tryptic and thermolysin peptides identified eight intramolecular disulfide bonds involving all 16 conserved cysteine residues within AMA-1. Structural analysis of peptides from those peaks identified as containing disulfide bonded peptides in the nonreducing chromatographic profile revealed no heterogeneity in the cysteine connectivity pattern. However, an ambiguity in the cysteine connectivities remains to be resolved: it is clear that Cys431 and Cys433 link with Cys448 and Cys450, but it is not clear whether 431 links with 448 and 433 with 450 or vice versa. This ambiguity is difficult to resolve, because these four cysteines are found in two Cys-Xaa-Cys sequences lacking obvious chemical or enzymatic cleavage sites.

The arrangement of the disulfide bonds (Fig. 5) suggests that the ectodomain of AMA-1 may be composed of three subdomains characterized by three, two, and three disulfide bonds with masses of approximately 19, 13, and 13 kDa, respectively. Searches of available protein and DNA sequence data bases have failed to identify other proteins with any significant sequence relationship to AMA-1. However, the putative subdomain structure for AMA-1 enables the disulfide bond architecture to be compared with that of other proteins. This has not been fruitful for the subdomains I and II, but although the three disulfide bonds in subdomain III have not been fully resolved, one of the two linkage patterns possible closely resembles the cystine knot motif found in growth factors (23, 24). In such structures two of the disulfide bonds are involved in localized cyclization of the sequence, via two Cys-Xaa-Cys motifs, whereas, the third disulfide bond passes through the ring structure. Also common to these cysteine knot structures are extended segments of twisted anti-parallel b-strand, which lie between the first to second and fourth to fifth cysteines within the motif (25). Structural analyses of a series of AMA-1 homologues, using PredictProtein Server (26–28), indicated a high probability of b-strand in analogous locations of the cysteine knot-like structure of AMA-1 (data not shown).

The distribution of mutations within the sequence of AMA-1 is not uniform. The shaded circles in Fig. 6a show the location of amino acid substitutions that occur among the AMA-1 sequences of the P. chabaudi adami strains DS and 556KA and P. chabaudi chabaudi strain CB. Each of the putative subdomains defined by the disulfide bond connectivities exhibits some sequence diversity; however, subdomain I is the most diverse containing 58% of all sites where amino acid substitutions occur in the ectodomain of P. chabaudi AMA-1. A second, smaller hypervariable region also exists around the eighth cysteine residue (Cys262) in subdomain II. The relationship between sites of amino acid substitution and the disulfide bonds differs among the subdomains. In subdomain I, the most diverse of the subdomains, the majority of the amino acid substitutions occur between the first and third cysteine residues with the disulfide bonded “core” of this subdomain relatively free of amino acid substitutions (Fig. 6a). However, in subdomains II and III (Fig. 6a) amino acid substitutions that are well separated from each other within the primary sequence, are close to cysteine residues, and in the folded protein will be clustered as a result of disulfide bond formation. A similar relationship between the pattern of mutations and disulfide bonds is apparent in P. falciparum AMA-1 (Fig. 6b) with the clustering of the amino acid substitutions in subdomains II and III, as a result of the disulfide bond formation, being more pronounced than that seen in P. chabaudi AMA-1. Thus, these amino acid substitutions may define conformational, disulfide bond-dependent epitopes that are recognized by protective immune responses.

Consistent with this conclusion, we have recently established that some of the disulfide bonded tryptic fragments are recognized by antibodies induced by P. chabaudi infection in mice. These antibodies react with refolded AMA-1 but not the reduced and alkylated antigen. We are currently producing the putative subdomains in an E. coli expression system and synthesizing disulfide bonded peptides based on the cysteine connectivity pattern within AMA-1 in an effort to establish which region(s) of the molecule are most relevant for the generation of a protective immune response.

Acknowledgments—We thank Mai Hayden, Stirling Edwards, and David Pye of CSL Ltd., Melbourne, Australia, who provided the culture supernatant containing recombinant AMA-1 that was used as starting material in this study. We are particularly grateful to Etty Bonnici for typing this manuscript.

REFERENCES

1. Anders, R. F., and Saul, A. J. (1993) Molecular Immunological Considerations in Malaria Vaccine Development (Good, M., and Saul, A. J., eds) pp. 169–208, CRC Press, London

2. Deans, J. A., Thomas, A. W., and Cohen, S. (1984) Mol. Biochem. Parasitol. 11, 189–204

3. Peterson, M. G., Marshall, V. M., Smythe, J. A., Crotwell, P. E., Lew, A. M., Silva, A., Anders, R. F., and Kemp, D. J. (1989) Mol. Cell. Biol. 9, 3151–3154

4. Marshall, V. M., Zhang, L. X., Anders, R. F., and Coppe!, R. L. (1996) Mol. Biochem. Parasitol. 77, 109–113

5. Petersen, M. G., Nguyen-Dinh, P., Marshall, V. M., Eliot, J. F., Collins, W. E., Anders, R. F., and Kemp, D. J. (1996) Mol. Biochem. Parasitol. 79, 279–284

6. Marshall, V. M., Petersen, M. G., Lew, A. M., and Kemp, D. J. (1989) Mol. Biochem. Parasitol. 37, 281–284

7. Cheng, Q., and Saul, A. J. (1994) Mol. Biochem. Parasitol. 65, 183–187

8. Dutta, S., Malhotra, P., and Chauhan, V. S. (1995) Mol. Biochem. Parasitol. 73, 267–270

9. Narum, D. L., and Thomas, A. W. (1994) Mol. Biochem. Parasitol. 67, 59–68

10. Deans, J. A., Knight, A. M., Jean, W. C., Waters, A. P., Cohen, S., and Mitchell, G. H. (1988) Parasite Immunol. 10, 535–552

11. Thomas, A. W., Deans, J. A., Mitchell, G. H., Alderson, T., and Cohen, S. (1984) Mol. Biochem. Parasitol. 13, 187–199

12. Collins, W. E., Pye, D., Crotwell, P. E., Van den Berg, K. L., Gaolland, G. G., Sulzer, A. J., Kemp, D. L., Edwards, S. J., Coppe!, R. L., Sullivan, J. S., Morris, C. L., and Anders, R. F. (1994) Am. J. Trop. Med. Hyg. 51, 711–719

13. Thomas, A. W., Waters, A. P., and Carr, D. (1990) Mol. Biochem. Parasitol. 42, 285–298

14. Waters, A. P., Thomas, A. W., Mitchell, G. H., and McCutchan, T. F. (1991) Mol. Biochem. Parasitol. 44, 143–144

15. Laemmli, U. K. (1970) Nature 227, 680–685

16. Tobin, H., Staehelin, T., and Gordon, J. (1979) Proc. Natl. Acad. Sci. U. S. A. 76, 4350–4352

17. Moritz, R. L., Eddes, J. J., Ji, H., Reid, G. E., and Simpson, R. J. (1995) Techniques in Protein Chemistry 6 (Crabb, J. W., ed) pp. 311–322, Academic Press, San Diego

18. Reid, G. E., Ji, H., Eddes, J. J., Moritz, R. L., and Simpson, R. J. (1995) Electrophoresis 16, 1120–1130

19. Moritz, R. L., Reid, G. E., Ward, L. D., and Simpson, R. J. (1994) Methods: A Companion to Methods in Enzymology 6, 213–220

20. Moritz, R. L., and Simpson, R. J. (1992) J. Chromatogr. 599, 119–130

21. Moritz, R. L., and Simpson, R. J. (1992) J. Microcolumn Sep. 5, 485–489

22. Krishna, R. G., and Wold, F. (1983) Methods in Protein Sequence Analysis (Imahori, K., and Sakiyama, F., eds) pp. 107–172, Plenum Publishing Corp., New York

23. Sun, P. D., and Davies, D. R. (1995) Annu. Rev. Biophys. Biomol. Structure 24, 289–291

24. Isaacs, N. W. (1995) Curr. Opin. Struct. Biol. 5, 391–395

25. McDonald, N. Q., and Hendrickson, W. A. (1993) Cell 73, 421–424

26. Rost, B., Sander, C., and Schneider, R. (1994) Comput. Appl. Biochem. 10, 53–60

27. Rost, B., and Sander, C. (1993) J. Mol. Biol. 232, 584–599

28. Rost, B., and Sander, C. (1994) Proteins 19, 55–72

* A. N. Hodder and R. F. Anders, manuscript in preparation.