Published in final edited form as:

Biochimie. 2010 October ; 92(10): 1287–1295. doi:10.1016/j.biochi.2010.06.001.

Identification of key residues involved in fibril formation by the conserved N-terminal region of *Plasmodium falciparum* merozoite surface protein 2 (MSP2)

Xiaodong Yang^a, Christopher G. Adda^b, Christopher A. MacRaid^a, Andrew Low^a, Xuecheng Zhang^{a,b}, Weiguang Zeng^c, David C. Jackson^c, Robin F. Anders^b, and Raymond S. Norton^a

^aThe Walter & Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville 3052, Australia

^bDepartment of Biochemistry, La Trobe University, Bundoora, 3083, Australia

^cDepartment of Microbiology & Immunology, The University of Melbourne, Parkville 3010, Australia

Abstract

Merozoite surface protein 2 (MSP2) from the human malaria parasite *Plasmodium falciparum* is expressed as a GPI-anchored protein on the merozoite surface. MSP2 is assumed to have a role in erythrocyte invasion and is a leading vaccine candidate. Recombinant MSP2 forms amyloid-like fibrils upon storage, as do peptides corresponding to sequences in the conserved N-terminal region, which constitutes the structural core of fibrils formed by full-length MSP2. We have investigated the roles of individual residues in fibril formation and local ordered structure in two peptides, a recombinant 25-residue peptide corresponding to the entire N-terminal domain of mature MSP2 and an 8-residue peptide from the central region of this domain (residues 8–15). Both peptides formed fibrils that were similar to amyloid-like fibrils formed by full-length MSP2. Phe11 and Ile12 have important roles both in stabilising local structure in these peptides and promoting fibril formation; the F11A and I12A mutants of MSP2_{8–15} were essentially unstructured in solution and fibril formation at pH 7.4 and 4.7 was markedly retarded. The T10A mutant showed intermediate behaviour, having a less well-defined structure than wild-type and slower fibril formation at pH 7.4. The mutation of Phe11 and Ile12 in MSP2_{1–25} significantly retarded but did not abolish fibril formation, indicating that these residues also play a key role in fibril formation by the entire N-terminal conserved region. These mutations had little effect on the aggregation of full-length MSP2, however, suggesting that regions outside the conserved N-terminus have unanticipated importance for fibril formation in the full-length protein.

Keywords

malaria; fibril; amyloid; nuclear magnetic resonance; structure

© 2010 Elsevier Masson SAS. All rights reserved

Correspondence Raymond S. Norton, The Walter & Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, 3050, Australia. Phone: +61 3 9345 2306. Fax: +61 3 9345 2686. ray.norton@wehi.edu.au.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Malaria remains a highly prevalent disease and a major cause of infant mortality in many of the world's poorest countries, although the increasingly widespread use of insecticide-impregnated bed-nets, artemisinin-combination therapy and other control measures has contributed to an improved malaria situation in some African countries [1,2]. The development of a vaccine against *Plasmodium falciparum*, the cause of the most serious form of human malaria, would provide a major new tool to complement existing malaria control measures. The most advanced vaccine, RTS,S [3], targets the pre-erythrocytic stages of the parasite but asexual blood-stage antigens, particularly merozoite surface proteins, are also considered important potential components of a malaria vaccine [4]. Data from a field trial in Papua New Guinea indicated that merozoite surface protein 2 (MSP2) is a particularly promising vaccine candidate. In this trial a vaccine containing a combination of three antigens, including the 3D7 allelic form of MSP2, significantly reduced parasite densities in 6–9 year-old children living in an area where the transmission intensity of *P. falciparum* is high [5].

MSP2 is a GPI-anchored membrane protein of ~ 30 kDa [6,7] and one of the most abundant proteins on the merozoite surface [8]. Although highly polymorphic, the nature of the central variable sequences in MSP2 allows all forms of the protein to be classified into either the 3D7 or FC27 allelic families [9,10]. The central variable region includes tandemly arrayed repeat sequences, the length, sequence and number of which differ between the different alleles [11]. The N- and C-terminal regions of MSP2 are highly conserved among all alleles. The N-terminal conserved region is 25 residues in length after cleavage of the predicted signal sequence whereas the C-terminal conserved region is ~50 residues in length after cleavage of the GPI-attachment signal (Figure 1).

Full-length MSP2 has the characteristics of an intrinsically unstructured protein [12], and recombinant proteins representative of both allelic types can form amyloid-like fibrils *in vitro* [13]. As each monomer is anchored into the merozoite membrane by the C-terminal GPI moiety, MSP2 cannot form fibrils *in situ* on the merozoite surface. However, MSP2 oligomers formed as a result of intermolecular β -strand interactions, similar to those responsible for amyloid formation by the recombinant protein, may contribute to the fibrillar projections comprising the merozoite surface coat [14]. The propensity of recombinant MSP2 to form fibrils under physiological conditions is consistent with this structural characteristic having relevance to the native protein but it is also an issue that needs to be addressed in developing MSP2 as a vaccine candidate. Defining the factors that govern fibril formation will facilitate the development of processes for producing either form of MSP2, and should enable the production of more stable vaccine formulations.

Peptides encompassing the conserved N-terminal region of MSP2 have been identified by proteinase K digestion as the core of the amyloid-like fibrils [13] and synthetic peptides from this region of the protein are capable of forming fibrils with amyloid-like properties [15]. One of these peptides, MSP2_{1–25}, corresponds to the entire N-terminal region that is conserved in all MSP2 alleles, commencing at the predicted N-terminus after cleavage of the signal peptide. The second is the central octapeptide SNTFINNA (MSP2_{8–15}), which was chosen because it corresponds to a region found to induce protective immunity in mice [16,17] and was identified as likely to form fibrils [15]. NMR analyses showed that MSP2_{8–15} formed a turn-like structure centred on Phe11 and Ile12 [15]; this was also present in MSP2_{1–25}, along with nascent helical structure in the C-terminal half of the peptide [18]. We have also characterized the backbone dynamics of MSP2_{1–25} in solution using peptide isotopically labelled in a bacterial expression system capable of producing milligram quantities of wild-type and mutant peptides [18]. In the present study we have investigated the roles of individual residues in both the solution conformation and propensity for fibril formation by synthetic and recombinant peptides from

the N-terminal region of MSP2. Our results identify features of MSP2 that are likely to influence its conformation and aggregation properties on the merozoite surface.

2. Materials and methods

2.1 Peptide synthesis

Peptides corresponding to MSP2₈₋₁₅ (Figure 1) and variants thereof were synthesized using Fmoc chemistry. The correct masses of all peptides were verified by mass spectrometry and the overall purities (>90%) were confirmed by reverse phase HPLC.

2.2 Expression and purification of MSP2₁₋₂₅ and full-length MSP2

PCR mutagenesis was used to construct the [F11A,I12A]MSP2₁₋₂₅ mutant with four 5' forward primers to amplify the entire MSP2₁₋₂₅. A final amplification cycle using a 5' forward primer was employed to engineer the thrombin site before the initiation codon (ATG) of [F11A,I12A]MSP2₁₋₂₅ mutant. The four 5' forward primers and thrombin site primer are as follows:

5' Primer I (FI-AA), TAGCAACACCGCTGCTAATAACGCCTATAA (30 bases)

5' Primer II (FI-AA) AAAGCAAATATAGCAACACCGCTGCTAATA (30 bases)

5' Primer III (FI-AA) ATTA AAAACGAAAGCAAATATAGCAACACC (30 bases)

5' Primer IV (FI-AA) CGGTTCTATGATTAAAAACGAAAGCAAATA (30 bases)

5' Thrombin Primer

CGGCGAATTCCTGGTGCCACGCGGTTCTATGATTAAAAACGAAAGC (46 bases)

A single 3' reverse primer, CGCCAAGCTTCTACGCCATCGAGCGGCG (28 bases), was used in all amplification cycles.

The amplification conditions, protein expression and purification procedures described previously for wild-type MSP2₁₋₂₅ and [Y7A,Y16A]MSP2₁₋₂₅ [18], were used to produce [F11A,I12A]MSP2₁₋₂₅. A shallow RP-HPLC gradient was employed to purify the mutant peptide, using 10 to 40 % acetonitrile (ChromAR HPLC grade, Mallinckrodt, Paris, KY) and 0.085 % (v/v) trifluoroacetic acid (HPLC/Spectro grade, Pierce, Rockford, IL) over 40 min. The purified [F11A,I12A]MSP2₁₋₂₅ peptide was lyophilised and stored at room temperature before NMR and fibril formation studies.

The molecular mass of [F11A,I12A]MSP2₁₋₂₅, was determined by MALDI-TOF mass spectrometry (Voyager Biospectrometry PR system; Applied Biosystems) with α -cyano-4-hydroxy-cinnamic acid (Sigma) matrix.

Full-length FC27 MSP2 carrying the mutations F11A, I12A and the double mutant [F11A,I12A] was prepared using Phusion site-directed mutagenesis (Finnzymes, Espoo, Finland), while 3D7 MSP2 carrying the same mutations was prepared using QuikChange site-directed mutagenesis (Stratagene, La Jolla, CA). Expression and purification of all full-length MSP2 mutants was performed as described previously [12].

2.3 NMR spectroscopy

NMR experiments on MSP2₈₋₁₅ and Ala-containing analogues were performed in 95% H₂O/5% ²H₂O and 10 mM acetic acid at pH 4.7 and 5 °C. Previously we have reported that the conformation of MSP2₈₋₁₅ was invariant over the pH range 3–7.3, as expected given that there are no titratable groups in the peptide with pK_a values in this range (the C-terminus was amidated) [15]. NMR data for structure calculations were therefore acquired at pH 4.7, where

all backbone amide resonances could be observed; it should be noted that aggregation and fibril formation do occur at this pH, albeit more slowly than at pH 7.4. NMR experiments on [F11A,I12A]MSP2_{1–25} were performed at a final concentration of 0.6 mM in 95% H₂O/5% ²H₂O containing 10 mM acetic acid at pH 3.4 and 5 °C. Two-dimensional homonuclear DQF-COSY, TOCSY (spin-lock time 37.5 ms) and NOESY (mixing times 100 and 250 ms) spectra were acquired on a DRX-600 spectrometer (Bruker Biospin). Water signals were suppressed using the WATERGATE pulse sequence [19]. A PFG longitudinal eddy-current delay pulse sequence was used for diffusion measurements [20–22]. All spectra were collected at 5 °C unless otherwise stated. Spectra were processed with TopSpin, version 1.3 (Bruker Biospin), and analyzed with XEASY, version 1.3.13 [23]. Chemical shifts for MSP2_{8–15} and MSP2_{1–25} have been deposited previously in the BMRB [24] with accession number 6698 and 15075, respectively. The chemical shifts of the various mutants of these peptides investigated in this study are tabulated in the Supplementary Material.

2.4 Structure calculations

Structures for MSP2_{8–15} mutants and [F11A,I12A]MSP2_{1–25} were calculated using distance restraints from NOESY spectra acquired at 600 MHz. Intensities of NOE cross-peaks measured in XEASY were calibrated using the CALIBA macro from the program CYANA [25]. Initial structures were calculated using torsion angle dynamics and simulated annealing protocols in CYANA and structures were optimized for a low target function. The same constraint set was also used to calculate a new family of 100 structures using XPLOR-NIH [26]. The 50 lowest energy structures were selected for energy minimization in a box of water. A final family of 20 lowest energy structures was chosen for analysis using PROCHECK-NMR [27] and MOLMOL [28]. The thirty lowest energy structures before water refinement had no experimental distance violations > 0.2 Å or dihedral angle violations > 5°. For the final 20 structures after water refinement, no distance violations > 0.2 or dihedral angle violations > 5° were present. In some cases the structures derived from CYANA were subjected directly to unrestrained energy minimization in a box of water using GROMACS [29]. Structural figures were prepared using MOLMOL [28] and PyMol (<http://www.pymol.org>).

2.5 Thioflavin T binding assays

Lyophilised peptide was reconstituted in Milli-Q grade water to a concentration of about 1 mg/ml, heated at 80 °C for 30 min and filtered through a 0.02 µm filter unit (Whatman) in order to remove preformed fibrils. Fibril formation reactions (1 ml) were established by diluting the peptide to 100 µM in PBS, pH 7.4; the samples were agitated during the assay. Full-length FC27 MSP2 was reconstituted in PBS to the final assay concentration (80 µM), then heated and filtered as above immediately prior to the start of the assay.

ThT binding assays were performed in 96-well microtitre plates (Nunc) in solutions containing 30 µM ThT and 100 µM peptide or 80 µM FC27 MSP2, and fibril formation was monitored by measuring ThT fluorescence (at excitation and emission wavelengths of 443 and 484 nm, respectively) using a SpectraMax M2 plate reader (Molecular Dynamics).

2.6 Size-exclusion chromatography assays

Full-length 3D7 MSP2 was reconstituted in PBS to 4 mg/ml and heated at 80 °C for 10 min and centrifuged at 15,000 g for 15 min at the start of the assay. Samples of 200 µl were incubated in sterile 1.5 ml tubes at room temperature for the course of the assay. 15 µl aliquots were taken and diluted to 1 mg/ml, centrifuged at 15,000 g for 15 min and further diluted to 300 µg/ml immediately prior to analysis on a Superdex 200 10/300 column (GE Healthcare Biosciences, Piscataway, NJ) in 10% acetonitrile in PBS at a flow rate of 0.4 ml/min.

2.7 Electron microscopy

Peptide and protein samples (10 μ L) were applied to 400 mesh copper grids coated with a thin layer of carbon. Excess material was removed by blotting and samples were negatively stained twice with 10 μ L of a 2 % (w/v) uranyl acetate solution (Electron Microscopy Services). The grids were air dried and viewed using a Jeol JEM-2010 transmission electron microscope operated at 80 kV. Peptide samples for EM were prepared within a week of the peptides being reconstituted in PBS for the dye-binding assays described above.

3. Results

3.1 Fibril formation by wild-type peptides

We have shown previously that recombinant MSP2₁₋₂₅ and synthetic MSP2₈₋₁₅ peptides formed amyloid-like fibrils rapidly in phosphate-buffered saline (PBS) at pH 7.4 [15,18]. The fibril formation assays for recombinant MSP2₁₋₂₅ and synthetic MSP2₈₋₁₅ peptides that are reported here have been further optimized. The optimized conditions require the peptides to be heated at 80 μ C prior to fibril formation in order to remove any aggregates that may seed fibril formation; the assays were then initiated with 100 μ M of peptide in PBS at pH 7.4 and 30 μ M of ThT. Fibril formation assays were performed at room temperature without agitation in 96-well microtitre plates using quadruplicate wells for each peptide. Under these conditions fibril formation by the recombinant 25-mer and synthetic 8-mer peptides showed similar kinetics, with a short lag phase followed by an exponential increase in fibril formation which reached a maximum in ~2.5 h (Figure 2). However, it should be noted that there was significant variability among replicate experiments in the length of the lag phase and the time taken to reach maximum fluorescence intensity.

The rate of fibril formation is highly pH-dependent. In acetate buffer without added NaCl at pH 3.4 and 4.7, MSP2₈₋₁₅ did not form ThT-reactive fibrils after > 30 days' incubation at room temperature. However, in NMR samples of wild-type and mutant MSP2₈₋₁₅, fibril formation occurred at these pH values over a period of months, as documented below. As the conformation of MSP2₈₋₁₅ in aqueous solution is not pH-dependent over the range 3–7 [15], we infer that fibril formation is inhibited by the presence of a positively charged N-terminal ammonium group at low pH as there are no titratable side chains in this peptide and the C-terminus is amidated. Alternatively, the presence of acetate ions at the two lower pH values may have retarded fibril formation.

3.2 Fibril formation by mutant peptides

The role of individual residues in fibril formation was investigated initially using a set of seven mutant MSP2₈₋₁₅ peptides in which each of the first seven residues was replaced by alanine. ThT binding was used to follow fibril assembly in PBS; this showed that replacement of residues with alanine could change both the rate of fibril assembly, by changing the length of the lag time, and the intensity of ThT fluorescence at which fibril assembly reached a plateau (Figure 3A). Replacement of the central hydrophobic residues had the most pronounced effect on fibril formation; no ThT binding was seen with either the F11A or I12A mutants, whereas fibril formation by the T10A mutant had an extended lag time. The other mutations had no, or only minor, effects on the lag time, but in all three mutants where Asn was replaced by Ala, maximum ThT fluorescence was diminished relative to wild-type. In contrast, replacing the hydroxyl-containing residues Ser8 and Thr10 by Ala resulted in enhanced ThT fluorescence (Figure 3A).

It is possible that these mutants exert their effect in this assay by altering the binding or fluorescence quantum yield of ThT, rather than by affecting fibril formation. To address this, we monitored MSP2₈₋₁₅ aggregation by NMR. Wild-type MSP2₈₋₁₅ showed a progressive

loss of signal intensity with time and eventually formed insoluble aggregates, which were visible in the NMR tube. Electron micrographs of samples removed from the NMR tube confirmed the presence of fibrils similar to those observed previously for this peptide [15]. This loss of intensity was specific to the peptide, as signals from the acetate buffer in these samples did not show any change. It apparently reflected self-association of the peptide, leading initially to soluble aggregates with line widths too broad to be visible in high-resolution NMR, and eventually to insoluble fibrils. The loss of signal intensity for wild-type MSP2_{8–15} and its mutants are plotted as a function of time in Figure 3B. The decay curves of most mutants are similar to wild-type, except for N13A, which did not aggregate to the same extent as the others, and F11A and I12A, which did not show any signal loss over the time period monitored. These results are consistent with the ThT assay and suggest that these mutations indeed inhibit fibril formation by MSP2_{8–15}.

Even though there was no loss of signal intensity for the F11A and I12A mutants and no evidence of fibril formation in the ThT-binding experiments described above, these mutant peptides were able to form fibrils eventually. When removed from the NMR tubes after more than a year, both samples bound ThT and contained fibrils (Figures 4 and S1). We cannot rule out that fibril formation was triggered in these samples by the mechanical disturbance associated with removal from the NMR tube. However, these observations indicate that, while fibril formation is significantly retarded in F11A and I12A, it is not completely abrogated.

The influence of Phe11 and Ile12 on fibril formation was explored further by mutating these residues in the recombinant MSP2_{1–25} peptide. In an earlier study we found that replacing both tyrosine residues in the sequence ([Y7A,Y16A]MSP2_{1–25}) did not affect the rate of fibril formation [18]. However, when both Phe11 and Ile12 of MSP2_{1–25} were mutated to Ala, the rate of fibril formation in PBS was significantly slower, with a longer lag time of 3.5 h, compared with 1 h for both wild-type MSP2_{1–25} and [Y7A,Y16A]MSP2_{1–25} mutant (Figure 5). Moreover, [F11A,I12A]MSP2_{1–25} took approximately 9 h to reach its maximum fluorescence intensity, as compared to 2 h for MSP2_{1–25} and [Y7A,Y16A]MSP2_{1–25}. Thus, mutations at Phe11 and Ile12 of MSP2_{1–25} retarded but did not completely inhibit fibril formation under physiological conditions. Fibril formation was also significantly inhibited under acidic conditions, as indicated by our translational diffusion data and the absence of any loss of NMR peak intensity (see below).

3.3 MSP2_{8–15} mutant structures

NMR experiments were performed on MSP2_{8–15} and its mutants in 10 mM acetic acid at pH 4.7 and 5 °C. Translational diffusion coefficients [22] of $2.1\text{--}2.2 \times 10^{-10} \text{ m}^2 \text{ s}^{-1}$ obtained for these peptides at 5 °C were in good agreement with previous values for MSP2_{8–15} [15] and indicated that all MSP2_{8–15} mutants were initially monomeric under these solution conditions.

Chemical shift assignments were completed for all backbone protons and the majority of side chain protons for all peptides at 5 °C (Tables S1–S8 in Supplementary Material). The deviations from random coil chemical shifts for C^αH and backbone amide resonances are plotted in Figure S2 (Supplementary Material). The differences between each of the mutants and wild-type MSP2_{8–15} are shown in Figure 6. The effects of each Ala substitution were mainly local, although more widespread changes were observed for T10A, F11A and I12A.

In the course of analyzing 2D NMR spectra to make resonance assignments it became apparent that the numbers of NOEs in spectra of the T10A mutant, and even more so the F11A and I12A mutants, were significantly lower than for wild-type or other mutants. This is shown graphically in Figure 7 and the actual numbers of NOEs are tabulated in the Supplementary Material (Tables S9 and S10).

As the dearth of NOEs for these mutants presumably reflects less ordered structures, we carried out structure calculations on all mutant peptides to ascertain whether this was the case. It is well known that defining the conformations of linear peptides lacking disulfide or other covalent constraints requires caution, as the presence of just a few medium- or long-range restraints can bias the calculated structures when the vast majority of restraints are from short-range interactions. Previously, we have calculated structures for MSP2₈₋₁₅ using CYANA, CYANA followed by unrestrained energy minimization in GROMACS, and CYANA followed by restrained simulated annealing and energy minimization in X-PLOR [15]. A similar set of calculations was undertaken for the S8A mutant, with similar results; the derived structures were influenced by the calculation procedure, indicating that they were not adequately defined by the NMR data (Figure S3), but they concur in showing a chain reversal in the vicinity of Phe11 and Ile12, even though the exact nature of the turn was not defined. With this caveat, we have calculated structures in CYANA for all of the MSP2₈₋₁₅ mutants. Structures for wild-type and two mutants with comparable numbers of NOEs (S8A and N9A) were similarly well defined, that of T10A was less well defined, reflecting the lower number of NOEs, and those for F11A and I12A were essentially undefined (Figure 8), consistent with the very low number of NOEs (Table S11 in the Supplementary Material summarizes the structural parameters and statistics for structures of all of the mutants calculated in CYANA). Thus, it appears that the Phe11 and Ile 12 side chains play key roles in stabilising local structure in these peptides, as well as in fibril formation.

3.4 [F11A,I12A]MSP2₁₋₂₅ structure

Wild-type MSP2₁₋₂₅ and [Y7A,Y16A]MSP2₁₋₂₅ both form fibrillar aggregates under the conditions used for our NMR studies (10 mM acetic acid at pH 3.4 and 5 °C), with visible aggregates formed in the NMR tube over a period of one month (data not shown). In contrast, no such aggregates were visible for [F11A,I12A]MSP2₁₋₂₅ and no loss of signal intensity was observed over a period of three months (Figure S4). The translational diffusion coefficients of [F11A,I12A]MSP2₁₋₂₅ at pH 3.4 and 5 °C, measured weekly over a period of 3 months, were $1.09\text{--}1.21 \times 10^{-10} \text{ m}^2 \text{ s}^{-1}$. These were essentially identical to its wild-type MSP2₁₋₂₅ counterpart and the [Y7A,Y16A]MSP2₁₋₂₅ mutant [18], suggesting that [F11A,I12A]MSP2₁₋₂₅ remained monomeric over 3 months under these conditions. This shows that mutations of Phe11 and Ile12 significantly retard the rate of fibril formation in acetate buffer, consistent with the results obtained in PBS with the ThT binding assay.

Sequence-specific chemical shift assignments for all backbone protons and the majority of side chain protons of [F11A,I12A]MSP2₁₋₂₅ are tabulated in the Supplementary Material (Table S12). The chemical shift deviations from random coil for C^αH and backbone amide resonances of [F11A,I12A]MSP2₁₋₂₅ at 5 °C are shown in Figure S5A. Backbone chemical shift differences between [F11A,I12A]MSP2₁₋₂₅ and wild-type MSP2₁₋₂₅ were mostly < 0.1 ppm in magnitude and localized to the flanking residues (Figure S5B). These differences were significantly less than those in [Y7A,Y16A]MSP2₁₋₂₅, which caused greater effects on residues flanking the double Tyr mutations as well as towards the C-terminus (Ile20 - Ser23) [18].

As for [S8A]MSP2₈₋₁₅, we calculated structures for [F11A,I12A]MSP2₁₋₂₅ using CYANA, CYANA-GROMACS and X-PLOR. The same conclusion applied, that they were not well defined by the NMR data (Figure S6). The NMR-derived restraints and structural statistics from X-PLOR for [F11A,I12A]MSP2₁₋₂₅ are summarized in Table S13. The angular order parameter plots (Figure S7) showed that ϕ angles were generally well-ordered ($S_\phi > 0.8$) but ψ angles were not. Fewer medium- and long-range NOEs were observed for [F11A,I12A]MSP2₁₋₂₅ (Figure S8A) than for wild-type MSP2₁₋₂₅ (Figure S8B), indicating a less ordered structure. In particular, the NOEs missing from the spectrum of the mutant were from the turn

and nascent helical structures in the central and C-terminal regions, indicating that these regions of structure were disrupted by the double mutation, consistent with our observations for MSP2₈₋₁₅.

3.5 Fibril formation by full-length MSP2

The N-terminal region of MSP2 constitutes the structural core of the fibrils formed by the full-length protein. For this reason, we hypothesized that the mutations that inhibit aggregation of the N-terminal peptides were likely to similarly inhibit aggregation of full-length MSP2. To test this we generated F11A and I12A single mutants, as well as the [F11A,I12A] double mutant, on the background of the FC27 and 3D7 allelic forms of MSP2. The kinetics of fibril formation of the mutant forms of FC27 MSP2 was compared to that of wild-type using ThT binding assays similar to those used for the peptides. The fibrils formed by 3D7 MSP2 do not bind ThT. Instead, the depletion of monomeric 3D7 MSP2 was monitored by size-exclusion chromatography (SEC). Surprisingly, we see no evidence that the mutations cause any inhibition of aggregation for full-length MSP2 of either the FC27 or 3D7 allelic form (Figure 9). We further examined the aggregates of the full-length 3D7 mutants by EM; the fibrillar morphology of these aggregates was identical to that of wild-type 3D7 MSP2 (Figure 10). Together, these data imply that fibril formation by full-length MSP2 of both allelic forms depends upon mechanisms distinct from those that drive fibril formation by the isolated N-terminal region.

4. Discussion

In this study, we have investigated the conserved N-terminal region of MSP2 and a shorter 8-residue sequence from this region with a particular focus on the roles of individual side chains in their solution conformations and their propensity for fibril formation. Both peptides contained turn-like structures centered on residues 10–12, with the longer peptide also showing turn-like or nascent helical structure towards the C-terminus, around residues 17–24. Our results show that the Phe11 and Ile12 side chains have important roles both in stabilising this local structure and promoting fibril formation in these peptides. The F11A and I12A mutants of the MSP2₈₋₁₅ peptide were essentially unstructured in solution (Figure 7) and fibril formation at pH 4.7 and 7.4 was drastically retarded (Figures 3 and S3), even though samples incubated for more than a year did form ThT-reactive fibrils. The T10A mutant showed intermediate behaviour, having a less defined structure than wild-type and slower fibril formation at pH 7.4. This strong correlation between local conformation in solution and propensity for fibril formation was surprising, as we expect that the turn-like structures seen in solution would be replaced by the cross- β structures typical of amyloid-like fibrils [13,30,31] as fibril formation proceeds. Presumably, the Phe11 and Ile12 side chains, and to a lesser extent that of Thr10, stabilize both of these two alternative structures, one in solution and the other in fibrils. As such, the effect of the Phe11 and Ile12 sidechains on both solution structure and aggregation propensity can most likely be understood as independent consequences of the hydrophobic nature of these residues.

The double mutation of Phe11 and Ile12 in [F11A,I12A]MSP2₁₋₂₅ significantly retarded but did not abolish fibril formation under physiological conditions, indicating that these residues also play a key role in fibril formation in this longer peptide. In addition, these mutations caused loss of local ordered structure in this peptide, as in MSP2₈₋₁₅. Several medium-range NOEs reflecting local ordered structure in the wild-type N-terminal peptide were missing from spectra of [F11A,I12A]MSP2₁₋₂₅. The local structure in the wild-type peptide consists of β -turn and nascent helical structures in the central and C-terminal regions [18], and it appears that both structural elements have been destabilized in [F11A,I12A]MSP2₁₋₂₅.

As these structural elements are governed principally by local interactions, we predict that they will also be present in full-length MSP2. Indeed, NMR analyses of full-length MSP2 found evidence of motional restriction and nascent helical structure in this region of the protein [12]. In addition, the N-terminal region of MSP2 forms the structural core of the full-length MSP2 fibrils, as shown by proteinase-K digestion of fibrils formed by both FC27 and 3D7 allelic forms of MSP2 [13]. For these reasons we expected mutation of Phe11 and Ile12 to have a similar effect on fibril formation by full-length MSP2 to that of the peptides. The observation that mutated full-length MSP2 formed fibrils with a similar propensity to the wild-type proteins suggests that the mechanism of aggregation of full-length MSP2, and perhaps details of the fibril structure, differs from that of the N-terminal peptides. Even in short peptides, sequence context plays a dramatic role in determining aggregation propensity [32]. In the case of MSP2, it is clear from the different rates of fibril formation by full-length FC27 and 3D7 that the variable region influences the rate of fibril formation, and we have shown that the fibrils formed in each case are not identical in their properties [13]. It appears, therefore, that, although the conserved N-terminal region constitutes the structural core of MSP2 fibrils, unidentified elements of the variable region alter the aggregation mechanism such that the side chains of Phe11 and Ile12 no longer play the critical role they do in fibril formation by the N-terminal peptides.

The propensity of recombinant MSP2 to form amyloid fibrils is an issue that needs to be addressed as this antigen is developed further as a potential component of a malaria vaccine [4]. The finding that fibril formation by the conserved N-terminal region of the molecule can be markedly delayed by just a few mutations suggests that it may be useful to incorporate these changes in MSP2 as part of a strategy to produce a form of this antigen that remains monomeric through the manufacturing process and during storage prior to formulation. Mutating this region of the molecule should not impair the efficacy of MSP2 as a vaccine component because the results of a vaccine trial in Papua New Guinea [5] indicate that MSP2-induced protective immune responses are directed against the central variable region of the protein, which appears to be preferentially displayed on the merozoite surface. However, our results also show that additional residues outside the N-terminal domain will need to be modified in order to block fibril formation by the full-length molecule. Alternatively, it may be desirable to remove the N-terminal region altogether from MSP2 in order to minimize aggregation whilst maintaining the immunologically important central region.

Because MSP2 readily forms amyloid-like fibrils under physiological conditions we have hypothesized that MSP2, which is one of the most abundant components of the merozoite surface coat, may form related polymers *in situ*, on the parasite surface. As the induction of protective antibody responses by vaccines is crucially dependent on the conformation of protein immunogens in vaccines it is important to establish if MSP2 on the merozoite surface is involved in homo- or hetero-oligomeric complexes and to what extent such interactions place constraints on the conformations adopted by this protein, which in its monomeric state is highly disordered [12].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Shenggen Yao, and Jennifer Sabo for valuable discussions and Brian Smith for assistance with GROMACS. This work was supported in part by grants from the Australian Research Council (DP0344826 and DP0664723) and the National Institutes of Health (NIH R01AI59229), as well as a National Health and Medical Research Council (NHMRC) IRIISS Grant 361646 and a Victorian State Government OIS grant. R.S.N. acknowledges fellowship support from the NHMRC.

Abbreviations

DQF-COSY	double quantum filtered correlation spectroscopy
EM	electron microscopy
MSP2	merozoite surface protein 2
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser and exchange spectroscopy
PBS	phosphate buffered saline
TOCSY	total correlation spectroscopy
r.m.s.d.	root mean square deviation
ThT	thioflavin T

References

- [1]. Greenwood B. Between hope and a hard place. *Nature* 2004;430:926–927. [PubMed: 15318230]
- [2]. Greenwood BM, Fidock DA, Kyle DE, Kappe SH, Alonso PL, Collins FH, Duffy PE. Malaria: progress, perils, and prospects for eradication. *The Journal of clinical investigation* 2008;118:1266–1276. [PubMed: 18382739]
- [3]. Aponte JJ, Aide P, Renom M, Mandomando I, Bassat Q, Sacarlal J, Manaca MN, Lafuente S, Barbosa A, Leach A, Lievens M, Vekemans J, Sigauque B, Dubois MC, Demoitie MA, Sillman M, Savarese B, McNeil JG, Macete E, Ballou WR, Cohen J, Alonso PL. Safety of the RTS,S/AS02D candidate malaria vaccine in infants living in a highly endemic area of Mozambique: a double blind randomised controlled phase I/IIb trial. *Lancet* 2007;370:1543–1551. [PubMed: 17949807]
- [4]. Anders RF, Adda CG, Foley M, Norton RS. Recombinant protein vaccines against the asexual blood-stages of *Plasmodium falciparum*. *Human Vaccines* 2010;6:39–53. [PubMed: 20061790]
- [5]. Genton B, Betuela I, Felger I, Al-Yaman F, Anders RF, Saul A, Rare L, Baisor M, Lorry K, Brown GV, Pye D, Irving DO, Smith TA, Beck HP, Alpers MP. A recombinant blood-stage malaria vaccine reduces *Plasmodium falciparum* density and exerts selective pressure on parasite populations in a phase 1-2b trial in Papua New Guinea. *J Infect Dis* 2002;185:820–827. [PubMed: 11920300]
- [6]. Gerold P, Schofield L, Blackman MJ, Holder AA, Schwarz RT. Structural analysis of the glycosylphosphatidylinositol membrane anchor of the merozoite surface proteins-1 and -2 of *Plasmodium falciparum*. *Mol. Biochem. Parasitol* 1996;75:131–143. [PubMed: 8992312]
- [7]. Smythe JA, Coppel RL, Brown GV, Ramasamy R, Kemp DJ, Anders RF. Identification of two integral membrane proteins of *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. U S A* 1988;85:5195–5199. [PubMed: 3293051]
- [8]. Gilson PR, Nebl T, Vukcevic D, Moritz RL, Sargeant T, Speed TP, Schofield L, Crabb BS. Identification and stoichiometry of glycosylphosphatidylinositol-anchored membrane proteins of the human malaria parasite *Plasmodium falciparum*. *Mol Cell Proteomics* 2006;5:1286–1299. [PubMed: 16603573]
- [9]. Fenton B, Clark JT, Wilson CF, McBride JS, Walliker D. Polymorphism of a 35–48 kDa *Plasmodium falciparum* merozoite surface antigen. *Mol. Biochem. Parasitol* 1989;34:79–86. [PubMed: 2651919]
- [10]. Smythe JA, Coppel RL, Day KP, Martin RK, Oduola AM, Kemp DJ, Anders RF. Structural diversity in the *Plasmodium falciparum* merozoite surface antigen 2. *Proc Natl Acad Sci U S A* 1991;88:1751–1755. [PubMed: 2000383]
- [11]. Franks S, Baton L, Tetteh K, Tongren E, Dewin D, Akanmori BD, Koram KA, Ranford-Cartwright L, Riley EM. Genetic diversity and antigenic polymorphism in *Plasmodium falciparum*: extensive serological cross-reactivity between allelic variants of merozoite surface protein 2. *Infect. Immun* 2003;71:3485–3495. [PubMed: 12761133]

- [12]. Zhang X, Perugini MA, Yao S, Adda CG, Murphy VJ, Low A, Anders RF, Norton RS. Solution conformation, backbone dynamics and lipid interactions of the intrinsically unstructured malaria surface protein MSP2. *Journal of molecular biology* 2008;379:105–121. [PubMed: 18440022]
- [13]. Adda CG, Murphy VJ, Sunde M, Waddington LJ, Culvenor JG, Vingas K, Anderson V, Masciantonio R, Howlett GJ, Hodder AN, Anders RF. *Plasmodium falciparum* merozoite surface protein 2 is unstructured and forms amyloid-like fibrils. *Mol. Biochem. Parasitol* 2009;166:159–171. [PubMed: 19450733]
- [14]. Langreth SG, Jensen JB, Reese RT, Trager W. Fine structure of human malaria *In Vitro*. *J. Protozol* 1978;25:443–452.
- [15]. Yang X, Adda CG, Keizer DW, Murphy VJ, Rizkalla MM, Perugini MA, Jackson DC, Anders RF, Norton RS. A partially structured region of a largely unstructured protein, *Plasmodium falciparum* merozoite surface protein 2 (MSP2), forms amyloid-like fibrils. *J Pept Sci* 2007;13:839–848. [PubMed: 17883245]
- [16]. Lougovskoi AA, Okoyeh NJ, Chauhan VS. Mice immunised with synthetic peptide from N-terminal conserved region of merozoite surface antigen-2 of human malaria parasite *Plasmodium falciparum* can control infection induced by *Plasmodium yoelii yoelii* 265BY strain. *Vaccine* 1999;18:920–930. [PubMed: 10580206]
- [17]. Saul A, Lord R, Jones GL, Spencer L. Protective immunization with invariant peptides of the *Plasmodium falciparum* antigen MSA2. *J Immunol* 1992;148:208–211. [PubMed: 1727867]
- [18]. Low A, Chandrashekar IR, Adda CG, Yao S, Sabo JK, Zhang X, Soetopo A, Anders RF, Norton RS. Merozoite surface protein 2 (MSP2) of *Plasmodium falciparum*: Expression, structure, dynamics and fibril formation of the conserved N-terminal domain. *Biopolymers* 2007;87:12–22. [PubMed: 17516503]
- [19]. Piotto M, Saudek V, Sklenář V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR* 1992;2:661–665. [PubMed: 1490109]
- [20]. Dingley AJ, Mackay JP, Chapman BE, Morris MB, Kuchel PW, Hambly BD, King GF. Measuring protein self-association using pulsed-field-gradient NMR spectroscopy: application to myosin light chain 2. *J. Biomol. NMR* 1995;6:321–328. [PubMed: 8520223]
- [21]. Gibbs SJ, Johnson CS. A PFG-NMR Experiment for Accurate Diffusion and Flow Studies in the Presence of Eddy Currents. *J. Magn. Reson* 1991;93:395–402.
- [22]. Yao S, Howlett GJ, Norton RS. Peptide self-association in aqueous trifluoroethanol monitored by pulsed field gradient NMR diffusion measurements. *J Biomol NMR* 2000;16:109–119. [PubMed: 10723990]
- [23]. Bartels C, Xia TH, Billeter M, Güntert P, Wüthrich K. The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J Biomol NMR* 1995;6:1–10.
- [24]. Seavey BR, Farr EA, Westler WM, Markley JL. A relational database for sequence-specific protein NMR data. *J Biomol NMR* 1991;1:217–236. [PubMed: 1841696]
- [25]. Herrmann T, Güntert P, Wüthrich K. Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *J Biomol NMR* 2002;24:171–189. [PubMed: 12522306]
- [26]. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM. The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* 2003;160:65–73. [PubMed: 12565051]
- [27]. Laskowski RA, Rullmann JA, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 1996;8:477–486. [PubMed: 9008363]
- [28]. Koradi R, Billeter M, Wüthrich K. MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* 1996;14:51–55. 29–32. [PubMed: 8744573]
- [29]. Lindahl E, Hess B, Van Der Spoel D. GROMACS: A package for molecular simulation and trajectory analysis. *J Mol Model* 2001;7:306–317.
- [30]. Nelson R, Sawaya MR, Balbirnie M, Madsen AO, Riekel C, Grothe R, Eisenberg D. Structure of the cross- β spine of amyloid-like fibrils. *Nature* 2005;435:773–778. [PubMed: 15944695]
- [31]. Sawaya MR, Sambashivan S, Nelson R, Ivanova MI, Sievers SA, Apostol MI, Thompson MJ, Balbirnie M, Wiltzius JJ, McFarlane HT, Madsen AO, Riekel C, Eisenberg D. Atomic structures

of amyloid cross-beta spines reveal varied steric zippers. *Nature* 2007;447:453–457. [PubMed: 17468747]

- [32]. Maurer-Stroh S, Debulpaep M, Kuemmerer N, Lopez de la Paz M, Martins IC, Reumers J, Morris KL, Copeland A, Serpell L, Serrano L, Schymkowitz JWH, Rousseau F. Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. *Nature Methods* 2010;7:237–242. [PubMed: 20154676]

Figure 1.

Sequences of peptides used in this study and the relationships between them. S: 19-residue signal peptide which is predicted to be cleaved to yield the mature protein. N: 25-residue N-terminal region conserved between allelic families. Variable region: Central repeat region that contains the highly polymorphic region of MSP2. C: Conserved 74-residue C-terminal domain with a disulfide bond. A: GPI Anchor. The MSP2₁₋₂₅ and MSP2₈₋₁₅ (boxed) peptides shown formed the basis of this study. All numbering of residues in this study is relative to mature full-length MSP2.

Figure 2.

Fibril formation by wild-type MSP₂₈₋₁₅ (●) and MSP₂₁₋₂₅ (□). The peptides were heated at 80 °C and reactions were established with 100 μM of peptide in PBS at pH 7.4 and incubated at room temperature. Fibril formation was monitored by measuring ThT fluorescence.

Figure 3.

Rates of fibril formation of wild-type MSP28-15 and seven mutants. (A) Reactions were established by diluting the peptide to 100 μ M in PBS, pH 7.4. At various time points, samples were diluted into PBS containing 30 μ M ThT and fibril formation was monitored by measuring ThT fluorescence and plotted against time. (B) Relative methyl ¹H NMR signal intensities as a function of time for MSP28-15 and mutants.

Figure 4. Electron micrographs of the fibrils formed by (A) MSP₂₈₋₁₅ N9A, (B) MSP₂₈₋₁₅ F11A and (C) MSP₂₈₋₁₅ I12A. Samples were taken from the NMR tubes (cf. Figure 3B), negatively stained and viewed using transmission EM. The scale bar represents 100 nm.

Figure 5. Kinetics of fibril formation by wild-type MSP2₁₋₂₅ (□), [Y7A,Y16A]MSP2₁₋₂₅ (●) and [F11A,I12A]MSP2₁₋₂₅ (▲). In 96 well plates, quadruplicate reactions were established by diluting the peptide to 100 μ M in PBS containing 30 μ M ThT and fibril formation was monitored by measuring ThT fluorescence. The average values for each protein were expressed as a fraction of the maximum value and plotted against time.

Figure 6.

Backbone chemical shift differences from wild-type for MSP28-15 mutants at 5 °C. (A) S8A (B) N9A (C) T10A (D) F11A (E) I12A (F) N13A (G) N14A. Chemical shifts values of wild-type and mutant peptides were corrected for the differences between random coil values before subtracting wild-type values from mutant values.

Figure 7. NOE restraints numbers for MSP₂₈₋₁₅ mutant peptides at 5 °C. NOE numbers are obtained from NOESY spectra (mixing time 250 ms; 600 MHz) recorded at peptide concentration of 5 mg/mL in 10 mM sodium acetate at pH 4.7.

Figure 8. Families of structures for wild-type MSP28-15, and T10A, F11A and I12A mutants calculated using CYANA at same orientation. Structures were superimposed over the backbone heavy atoms (N, C $^{\alpha}$ and C) of residues Asn9 to Asn14.

Figure 9.

Kinetics of fibril formation of full-length MSP2 mutants. (A) Reactions were established with 40 μ M monomeric MSP2 for each of wt FC27 MSP2 (●), F11A FC27 MSP2 (○), I12A FC27 MSP2 (□), and [F11A,I12A] FC27 MSP2 (◇) in PBS containing 30 μ M ThT. The measured ThT fluorescence intensity was plotted against time. (B) Following the establishment of the fibril forming reactions with 160 μ M of wt 3D7 MSP2 (●), F11A 3D7 MSP2 (○), I12A 3D7 MSP2 (□), and [F11A,I12A] 3D7 MSP2 (◇), SEC was used to determine the amount of monomeric MSP2 remaining, represented by the area under the peak corresponding to monomeric MSP2, and was plotted for each time interval.

Figure 10.

Electron micrographs of the fibrils formed by (A) wt 3D7 MSP2, (B) F11A 3D7 MSP2 (C) I12A 3D7 MSP2 and (D) [F11A,I12A] 3D7 MSP2. Peptide conditions are same as those in Figure 9B. Samples were negatively stained and viewed using transmission EM. The scale bar represents 500 nm.