The merozoite surface protein 1 (MSP-1) is the most abundant protein on the surface of the erythrocyte-invading Plasmodium merozoite, the causative agent of malaria. MSP-1 is essential for merozoite formation, entry into and escape from erythrocytes, and is a promising vaccine candidate. Here, we present monomeric and dimeric structures of full-length MSP-1. MSP-1 adopts an unusual fold with a large central cavity. Its fold includes several coiled-coils and shows structural homology to proteins associated with membrane and cytoskeleton interactions. MSP-1 formed dimers through the presence of the erythrocyte cytoskeleton protein spectrin, which may compete for the dimerization interface. Our work provides structural insights into the possible mode of interaction of MSP-1 with erythrocytes and establishes a framework for future investigations into the role of MSP-1 in Plasmodium infection and immunity.

INTRODUCTION

With 229 million reported cases and 409,000 deaths in 2019 worldwide and more than half of the world’s population at risk, malaria remains a great burden to global health (1). Earlier reductions in incident rates have leveled off between 2015 and 2019 (1), making the development of a vaccine as vital as ever. Merozoites are the erythrocyte-infecting form of the malaria-causing Plasmodium parasite, and merozoite surface proteins (MSPs) show potential as malaria vaccines [reviewed by Beeson et al. (2)]. After erythrocyte invasion, merozoites replicate within the parasitophorous vacuole (PV) for ~48 hours in the case of Plasmodium falciparum, until they egress by sequential rupture of the PV membrane (PVM) and host cell membrane and go on to invade new host cells within minutes (3). In this short time window, the merozoite surface is exposed to the immune system. Several integral and peripheral MSPs (4, 5) form a thick fibrillar coat on the merozoite surface, MSP-1 being the most abundant (5). Antibodies targeting MSP-1 are found in sera of individuals from malaria-endemic regions and have been shown to confer some immunity (6–8). Thus, MSP-1 is considered a prime candidate antigen for vaccine development (9). It consists of four subunits, p83, p30, p38, and p42, which are held together noncovalently (Fig. 1A). A glycosylphosphatidylinositol (GPI) anchor links the complex to the membrane via the p42 subunit. MSP-1 is produced by subtilisin-like protease 1 (SUB-1) cleavage of a ~190-kDa precursor protein (p190) just before erythrocyte egress. In a second processing step by SUB-2 during erythrocyte invasion, p42 is cleaved into p33 and p19; the latter remains attached to the parasite surface via its GPI anchor, while the rest of the complex is shed (10). The MSP-1 complex is known to interact with a range of other merozoite proteins, including MSP-3, MSP-6, MSP-7, and MSP-9, and MSP Duffy binding-like (MSPDBL)–1 and MSPDBL-2 (11–13), forming various different complexes on the parasite surface (14–16). MSP-7 associates with MSP-1 immediately after translation (17), while other MSPs bind to the complex after maturation (11, 18, 19). Although the functional role of MSP-1 is not yet well understood, it is essential to Plasmodium development given that MSP-1 knockout mutants could not be generated (20, 21). It is thought to play a role in early erythrocyte attachment and invasion (22); MSP-1 has been shown to bind the erythrocyte surface proteins glycophorin A (GPA) (23) and band 3 (15), the latter being essential in the merozoite invasion process (23). Other studies failed to detect direct binding of red blood cells (RBCs) by MSP-1 itself but reported interactions with other components of the MSP-1 complex, specifically MSPDBL-1, MSPDBL-2, MSP-6, and MSP-9, although the RBC host receptors of these components remain unknown, and conflicting data were reported for MSP-6 (12–15). MSP-1 also interacts with the erythrocyte cytoskeleton protein spectrin (24, 25) and plays a role in egress (25). Antibodies targeting all MSP-1 subunits (14, 26), as well as some associated MSPs (11, 13), have been shown to inhibit parasite growth in RBC cultures. Antibodies targeting the p19 subunit have received particular scrutiny (10, 20), although contradictory reports exist regarding their ability to prevent disease (6), and immunization trials using only p42 were unsuccessful in inducing significant protection against malaria (27). Given the additive inhibitory effect that was observed for antibodies targeting the different MSP-1 subunits in vitro parasite growth assays (26), full-length MSP-1 may be a better vaccine antigen. A recent phase 1a clinical trial reported promising immunogenicity of a full-length recombinant MSP-1 “heterodimer” composed of two peptide chains corresponding to the N- and C-terminal “halves” of MSP-1, p83/30 and p38/42, respectively (hdMSP-1; Fig. 1A) (9), which was previously shown to assemble in the same manner as the native MSP-1 complex (28).
Despite its clinical relevance, structural information on MSP-1 is sparse and limited to the C-terminal epidermal growth factor (EGF)–like domain consisting of subunit p19 (29, 30) and global mapping of intersubunit interactions (28). As expected from secondary structure prediction, circular dichroism reports a largely α-helical structure with a significant portion of unstructured regions (25), which may explain the difficulty in obtaining well-diffracting crystals of MSP-1. More detailed knowledge about the structure of MSP-1 may aid in elucidating its essential role in Plasmodium infectivity.

**RESULTS**

**Global architecture of MSP-1**

Using single-particle cryo–electron microscopy (cryo-EM), we solved the structures of full-length MSP-1 from *P. falciparum*, the deadliest and most prevalent of the *Plasmodium* spp. infecting humans (1). MSP-1 was produced recombinantly as a heterodimer (hdMSP-1; Fig. 1A) in the absence of any of the other natively associated MSPs, as used in the aforementioned successful phase 1a clinical trial (9), using the previously described MSP-1D construct based on the *P. falciparum* strain 3D7 (28). Structural analysis was performed for both hdMSP-1 and its fully SUB-1–processed form (henceforth designated MSP-1; Fig. 1A and fig. S1). The data showed continuous flexibility (movie S1) and yielded several subtly different conformations (fig. S2 and table S1). Specifically, we solved six monomeric structures from the MSP-1 dataset, at nominal resolutions ranging from 3.1 to 3.6 Å and two dimeric structures from the hdMSP-1 dataset at nominal resolutions of 3.3 and 3.6 Å, which were used to build atomic models de novo (figs. S2 to S4). The highest-resolution monomeric structure was designated as the main conformation and is used for further analysis here.

MSP-1 has no sequence orthologs outside of *Plasmodium* and adopts an unusual three-dimensional (3D) architecture with a central cavity, which is likely to face away from the parasite surface (Fig. 1B). The monomer consists of 37 α helices (Fig. 1C), which can be roughly divided into two interconnected domains found on either side of the cavity (Fig. 1, B and C). On one side, p38 and p42 form a helical bundle (H30 to H37) which were used to build atomic models de novo (figs. S2 to S4). The highest-resolution monomeric structure was designated as the main conformation and is used for further analysis here.

MSP-1 has no sequence orthologs outside of *Plasmodium* and adopts an unusual three-dimensional (3D) architecture with a central cavity, which is likely to face away from the parasite surface (Fig. 1B). The monomer consists of 37 α helices (Fig. 1C), which can be roughly divided into two interconnected domains found on either side of the cavity (Fig. 1, B and C). On one side, p38 and p42 form a helical bundle (H30 to H37) which were used to build atomic models de novo (figs. S2 to S4). The highest-resolution monomeric structure was designated as the main conformation and is used for further analysis here.

MSP-1 has no sequence orthologs outside of *Plasmodium* and adopts an unusual three-dimensional (3D) architecture with a central cavity, which is likely to face away from the parasite surface (Fig. 1B). The monomer consists of 37 α helices (Fig. 1C), which can be roughly divided into two interconnected domains found on either side of the cavity (Fig. 1, B and C). On one side, p38 and p42 form a helical bundle (H30 to H37) which were used to build atomic models de novo (figs. S2 to S4). The highest-resolution monomeric structure was designated as the main conformation and is used for further analysis here.

**RESULTS**

**Global architecture of MSP-1**

Using single-particle cryo–electron microscopy (cryo-EM), we solved the structures of full-length MSP-1 from *P. falciparum*, the deadliest and most prevalent of the *Plasmodium* spp. infecting humans (1). MSP-1 was produced recombinantly as a heterodimer (hdMSP-1; Fig. 1A) in the absence of any of the other natively associated MSPs, as used in the aforementioned successful phase 1a clinical trial (9), using the previously described MSP-1D construct based on the *P. falciparum* strain 3D7 (28). Structural analysis was performed for both hdMSP-1 and its fully SUB-1–processed form (henceforth designated MSP-1; Fig. 1A and fig. S1). The data showed continuous flexibility (movie S1) and yielded several subtly different conformations (fig. S2 and table S1). Specifically, we solved six monomeric structures from the MSP-1 dataset, at nominal resolutions ranging from 3.1 to 3.6 Å and two dimeric structures from the hdMSP-1 dataset at nominal resolutions of 3.3 and 3.6 Å, which were used to build atomic models de novo (figs. S2 to S4). The highest-resolution monomeric structure was designated as the main conformation and is used for further analysis here.

MSP-1 has no sequence orthologs outside of *Plasmodium* and adopts an unusual three-dimensional (3D) architecture with a central cavity, which is likely to face away from the parasite surface (Fig. 1B). The monomer consists of 37 α helices (Fig. 1C), which can be roughly divided into two interconnected domains found on either side of the cavity (Fig. 1, B and C). On one side, p38 and p42 form a helical bundle (H30 to H37) which were used to build atomic models de novo (figs. S2 to S4). The highest-resolution monomeric structure was designated as the main conformation and is used for further analysis here.

MSP-1 has no sequence orthologs outside of *Plasmodium* and adopts an unusual three-dimensional (3D) architecture with a central cavity, which is likely to face away from the parasite surface (Fig. 1B). The monomer consists of 37 α helices (Fig. 1C), which can be roughly divided into two interconnected domains found on either side of the cavity (Fig. 1, B and C). On one side, p38 and p42 form a helical bundle (H30 to H37) which were used to build atomic models de novo (figs. S2 to S4). The highest-resolution monomeric structure was designated as the main conformation and is used for further analysis here.

MSP-1 has no sequence orthologs outside of *Plasmodium* and adopts an unusual three-dimensional (3D) architecture with a central cavity, which is likely to face away from the parasite surface (Fig. 1B). The monomer consists of 37 α helices (Fig. 1C), which can be roughly divided into two interconnected domains found on either side of the cavity (Fig. 1, B and C). On one side, p38 and p42 form a helical bundle (H30 to H37) which were used to build atomic models de novo (figs. S2 to S4). The highest-resolution monomeric structure was designated as the main conformation and is used for further analysis here.
and cytoskeleton interactions (fig. S5). A flexible domain, which we term the wing domain, made up of the shorter N-terminal α helices of p83 (H1 to H12) is found on the opposite side of the central cavity (Figs. 1, B and C, and 2A). The EGF-like domain that makes up the C-terminal part of p42 (p19) was not resolved and is probably connected to the rest of the protein via a flexible linker including a region with no predicted secondary structure (Fig. 1, B and C, and figs. S6 and S7). The wing domain and the coiled-coil domain are held together via intersubunit interactions mediated by long helices extending outward from each domain and p30. The largest inter-subunit contact surface is found between p30 and p38 (2900 ± 100 Å², average of all MSP-1 conformations). The 67–amino acid–long H30 of p38 extends from the coiled–coil domain on one end of the central cavity to the opposite end, where it makes extensive contacts with H24 of p30, stabilized further by interactions with helices H26 to H29 that wrap around the H24/30 helix pair (Fig. 2). In addition, a long left-handed coiled coil made up of H14 to H18 of p83 kinks around the H24/30 pair and extends from the edge of the wing domain to the bottom end of the coiled–coil domain on the opposite side of cavity where H17 and H18 contact p42 at the H35–36 loop and H37 (Fig. 2; total buried surface 970 ± 70 Å²).

Conformational flexibility
The highest degree of conformational flexibility was observed in the wing domain, and the solved conformations differed mainly in the position of this domain relative to the coiled-coil domain (Fig. 3, fig. S2, and movie S1). Consequently, the local resolution of the maps was lowest in this domain (fig. S3, E and F), and local refinement of the wing domain was performed to improve the map and aid model building of this region (figs. S2A, S3G, and S4A). In addition, cross-links between peptides from the wing domain observed by cross-linking mass spectrometry (XL-MS) were consistent with the model (fig. S4D). Nevertheless, side-chain density remained weak in certain regions (in particular, for H10 and H11), and the residual uncertainty in the model in this region is reflected in the relatively high local model B-factors (fig. S4C).

Because of the continuous nature of the intrinsic flexibility of MSP-1, classification of the data into a limited number of classes led to slightly different resolved conformations for the MSP-1 and hdMSP-1 wing domains. No further notable differences were observed between the protomers of the dimeric hdMSP-1 and the monomeric MSP-1 forms. The only exception is found in the pocket between H30 and H35 (fig. S8A). Here, we observe density in the hdMSP-1 maps that most likely corresponds to the loop between H34 and H35 containing the p38/p42 SUB-1 processing site, although the density is not clear enough to allow unambiguous model building (Fig. 1B and fig. S8A). No clear density was observed in this pocket for the SUB-1–treated MSP-1 sample where this loop was cleaved, and thus likely to be less restrained than in the hdMSP-1 sample where it is not cleaved (Fig. 1B and fig. S8A). The p38/p42 processing has been shown to be critical for parasite viability (25). Thus, the absence of global (resolved) structural changes between the hdMSP-1 and fully processed MSP-1 samples suggests that the functional importance of processing of the H34–35 loop

Fig. 2. Atomic model of MSP-1. (A) Atomic model built for the main conformation of MSP-1. Different subunits are colored in the same manner as in Fig. 1. (B) Examples of intersubunit interactions holding the complex together. Hydrogen bonds and salt bridges are indicated by dotted lines.
Conserved and variable regions of MSP-1

Because the MSP-1 sequence varies among *P. falciparum* strains, the question arises whether the structure of MSP-1 from other strains would differ from the structures presented here. The MSP-1 sequence can be divided into 17 blocks based on sequence variability as first described by Tanabe and co-workers (32) (Fig. 4A and fig. S6). They noted that the primary structure of *P. falciparum* MSP-1 has a dimorphic nature, and isoforms from different *P. falciparum* strains were broadly classified as belonging to one of two variants, termed K1 and MAD20 based on two representative isolates (32). Highest sequence homology was found for blocks 1 (comprising H1), 3 (H2 to H10), 5 (H11 and H12), 12 (H30 and H31), and 17 (the EGF-like domain; Fig. 4A). A similar pattern of MSP-1 sequence conservation is observed for MSP-1 isoforms from different *Plasmodium* spp. (Fig. 4B and fig. S7). The flexible p83 wing domain consists of a conserved core, composed of helices H5 to H7, H9, and H12 (Fig. 4, A and B). The core of the coiled-coil domain, made up of helices H31 to H33 and H35, also shows relatively high sequence homology across *P. falciparum* strains and *Plasmodium* spp. (Fig. 4, A and B). The most polymorphic region (block 2, 65 amino acids in MSP-1D) is located between H1 and H2 in the wing domain, is predicted to be unstructured (fig. S6), and was not resolved in the EM maps (Fig. 4A). Likewise, the second oligomorphic region (block 4) contains a predicted unstructured stretch of 16 amino acids between H10 and H11 at the edge of the wing domain that was also not resolved (Fig. 4A and figs. S6 and S7). This is echoed in blocks 6, 8, 10, 14, and 16, where unresolved (predicted unstructured) loops corresponded to the most variable regions (Fig. 4A and figs. S6 and S7). While the resolved parts of blocks 6 (H13 to H17), 10 (H26 to H30), and 16 (H35 to H37) also contain dimorphic regions, they show comparatively higher homology between different *P. falciparum* isolates and other *Plasmodium* spp. relative to the loop regions (Fig. 4A and figs. S6 and S7). Together, these observations suggest that the highest sequence variability is found in the flexible loop regions and that the global architecture of MSP-1 observed here may be conserved between different genotypes.

Monomer-dimer equilibrium

While MSP-1 is often depicted as monomeric (13, 25), dimeric forms of MSP-1 have previously been detected in detergent-resistant membranes of *P. falciparum* 3D7 schizonts by native polyacrylamide gel electrophoresis (PAGE) and chemical cross-linking (33). Here, both monomeric and dimeric forms were observed for MSP-1 and hdMSP-1, with 2D classification of different (control) EM datasets yielding varying fractions of dimer and monomer classes (fig. S8, B and C). The hdMSP-1 datasets used for high-resolution 3D reconstruction yielded a larger fraction of dimer classes (53 to 100% dimer; fig. S8C) compared to the dataset used for the SUB-1–processed MSP-1, which appeared to predominantly yield monomer classes (only 4% dimer; fig. S8B), supported by the apparent molecular weight of the samples used, gauged by size-exclusion chromatography (SEC; fig. S1A). Note that a smaller control dataset of a second MSP-1 sample yielded predominantly dimeric 2D class averages (89% dimer; fig. S8B) and a corresponding higher molecular weight estimate by SEC (fig. S1A), suggesting that SUB–processed MSP-1 is not monomeric per se. Microscale thermophoresis (MST) showed that in solution dimerization is concentration dependent with a dissociation constant (*K*<sub>D</sub>) of 60 ± 10 and 30 ± 20 nM for hdMSP-1 and MSP-1, respectively (Fig. 5A), suggesting that there is no significant difference in the dimerization propensity of the two MSP-1 forms, and the difference in monomer-to-dimer ratio observed in the EM data may have been introduced during sample or grid preparation. Given the observed low *K*<sub>D</sub> and high local concentration of MSP-1 on the merozoite surface, MSP-1 can be expected to form dimers on the parasite surface, as was observed for MSP-1 from *P. falciparum* 3D7 (33) on which the construct used here is based.

Whether or not MSP-1 forms dimers on the parasite surface is of interest because this could potentially shield the dimerization interface to some extent from immune detection. Dimerization is mediated mainly through interactions between p83 and p42, contributing
Antibodies directed against a region encompassing H11 to H16, which includes the dimerization interface H14, are associated with protection against Plasmodium falciparum malaria in adolescents from Mali (34). Furthermore, in the recent vaccine trial by Blank et al. (9) using hMDSP-1, epitopes were identified on H14, the H27-28 loop, and H37, which elicited immunoglobulin G (IgG) antibody responses putatively associated with inhibiting parasite growth in vitro (Fig. 7 and fig. S9). It is possible that these antibodies would destabilize the dimer, although whether this underlies their protective action remains an open question.

**MSP-1 domains involved in erythrocyte interactions and immunogenicity**

As discussed in the introduction, MSP-1 has been shown to interact with erythrocyte components (15, 23–25), as well as a range of other merozoite proteins (11–13, 15, 17), forming various different complexes on the parasite surface (14–16). The molecular details of these interactions remain unclear. Given that the core of the flexible wing domain does not participate in the stabilization of the MSP-1 complex itself, its relatively high degree of conservation suggests that this domain may play a role in these interactions, as previously proposed by Kauth et al. (28). In support of this hypothesis, antibodies against the p83 subunit that makes up this domain were particularly effective in inhibiting infection of erythrocytes by malaria parasites in in vitro assays (26). The p83 subunit was furthermore identified as a major target of opsonizing antibodies in sera from semi-immune individuals from Burkina Faso (35). While p83 is indeed also known to bind other MSPs [such as MSP-7, MSPDBL-1, and MSPDBL-2 (11, 13, 17)], an antibody targeting p83 inhibited parasite growth without affecting complex formation with these MSPs (14), suggesting that p83 is important for infection not only as a binding site for other MSPs but also, in particular, through its direct interaction with erythrocytes. Notably, residues 232 to 278 (in MSP-1D numbering) of the p83 subunit, encompassing H6 to H8 (Fig. 8A), were shown to interact with a Glu-rich extracellular region of GPA on the erythrocyte surface (23). The positively charged patch found at H6 to H8 suggests that the interaction might
be electrostatic (Fig. 6). Deletion of GPA reduced but did not completely abrogate parasitemia (23). Blank and co-workers (9) identified two conserved epitopes on H8, which elicited an IgG antibody response putatively associated with inhibiting parasite growth in vitro (Fig. 7).

Other MSP-1 subunits have also been implicated in interactions with RBCs. Residues 1052 to 1166 (in MSP-1D numbering) of the p38 subunit, encompassing H31 and part of H30 and H32 (Fig. 6), were shown to specifically bind erythrocytes via an unknown mechanism independent of GPA and sialic acid, a negatively charged sugar moiety abundant on the erythrocyte surface (36). While the RBC target remained unclear (36), strongly basic patches present on H30 to H32 near the bottom of the MSP-1 cavity (Fig. 6) suggest that long-range electrostatic interactions with other erythrocyte components could play a role. When recombinantly produced, this part of the p38 domain was recognized by sera of semi-immune individuals (36). Notably, the recombinant p38 fragment and polyclonal antibodies against it could inhibit parasite invasion of RBCs in vitro by nearly 70 and 90%, respectively (36), suggesting that it plays an important role in *Plasmodium* infectivity. Furthermore, semiconserved B cell epitopes putatively associated with parasite growth inhibition in vitro were also identified on H30 and H32 in the study by Blank et al. (9) (Fig. 7). The p38 subunit, as well as the p42 subunit, was further shown to interact with specific peptides of the erythrocyte anion transporter band 3 (37). These band 3 peptides appeared to have higher affinity for p42 and, in particular, for the p19 domain than for p38 (37). Considering that band 3 is the most abundant protein on the erythrocyte surface and essential in the merozoite invasion process, these interactions may be highly important for parasite attachment, although a molecular mechanism or precise MSP-1 binding site remains unclear (37).

Besides its interaction with band 3 (37), the p42 subunit has been reported to bind highly negatively charged heparin-like polysaccharides, which are abundant on the erythrocyte surface (22). The resolved part of the p42 subunit (p33) contains two strongly basic patches on H35 and at the bottom of H36 to H37 (Fig. 6), which may play a role in heparin binding; the latter patch coincides with a semiconserved B cell epitope putatively associated with parasite growth inhibition (Fig. 7), supporting its functional importance (9).

In addition to these erythrocyte surface targets important for parasite invasion, MSP-1 has also been shown to play a role in egress through interaction with the erythrocyte cytoskeleton protein spectrin (24, 25). A previous study using a recombinant fragment of p83 (encompassing H2 to H10) identified residues 291 to 344 (in MSP-1D numbering), corresponding to H9 and H10 (Fig. 8A) as the binding site for spectrin (24). However, interactions with other subunits were not investigated (24), and the highly negatively charged spectrin could be expected to also interact with positively charged patches present elsewhere on the MSP-1 complex. As stated above, the resolved parts of p42 and p38 appear to adopt a BAR/IMD-like or STAT-like t-snare fold (Fig. S5), which is associated with membrane and cytoskeleton interactions, and the most extensive positively charged surface region of MSP-1 is found on this domain (Fig. 6). The recombinant (hd)MSP-1 used here also bound spectrin (Fig. 8B), and in the presence of spectrin, hdMSP-1 appeared to dimerize less readily than in its absence (1.7 ± 0.9 μM versus 80 ± 20 nM, respectively; Fig. 8C), suggesting that spectrin may compete for the dimerization interface. This interface includes not only H37 of the p42 BAR/IMD-like coiled-coil domain but also H14 (p83), which is likewise part of a coiled-coil bundle (Fig. S5) and contains a positively charged region (Fig. 6). MSP-1 has been linked to PVM rupture in the parasite egress pathway (24, 25, 38). Immuno-EM previously identified MSP-1 in high abundance on the PVM in regions where the parasite and RBC membrane were apparently in contact (24). Furthermore, inhibition of SUB-1 has been shown to prevent the MSP-1–spectrin interaction (25) and PVM rupture (38). Thus, the MSP-1–spectrin interaction may play a role in PVM rupture, and given their fold, it is possible that the coiled-coil regions
**Fig. 6. Electrostatic surface potential.** MSP-1 model colored by electrostatic surface potential, where blue and red correspond to positive and negatively charged regions, respectively. Insets show same view of the model in cylinder representation for reference.

**Fig. 7. Immunogenic epitopes.** Immunogenic B cell epitopes identified in an MSP-1 vaccine trial (9) are mapped onto MSP-1. On the basis of correlation between IgG responses and parasite growth inhibition assay (GIA) activity of the various investigated sera, Blank et al. (9) identified epitopes putatively associated with GIA activity. Blue, no role; red, possible role in inhibiting parasite growth. Colored ovals indicate the degree of sequence conservation for each epitope: Dark purple, light purple, white, and green correspond to conserved, semiconserved (>50% conserved residues, rest dimorphic), mostly dimorphic (<50% conserved residues, rest dimorphic), and nonconserved epitopes. Dotted lines represent unresolved loops (scaled 1/3.5 relative to ribbon). More details can be found in fig. S9. aa, amino acids.
of MSP-1 are involved in this mechanism. Note that Das and co-workers (25) previously reported that cleavage of the p38/p42 processing site was necessary for binding of MSP-1 to inside-out erythrocyte ghost vesicles, which was attributed to interactions with spectrin, while we observe binding of spectrin to hdMSP-1 in which this processing site is not cleaved. The reason for this discrepancy could lie in experimental differences but remains unclear.

While we can only speculate about the molecular mechanism of these interactions, the flexibility of MSP-1 and, in particular, its wing domain may be central to its role as a major interaction surface for these various erythrocyte proteins as well as other MSPs. That the previously reported B cell epitopes putatively associated with in vitro parasite growth inhibition can be mapped to various regions of the 3D structure (Fig. 7) suggests a multifaceted mode of interaction. As previously discussed, this ability of MSP-1 to bind multiple erythrocyte targets via seemingly different and redundant mechanisms might enhance the efficiency of erythrocyte attachment and allow evasion of host cell immune responses (14, 37).

DISCUSSION

In summary, this work defines the architecture of MSP-1, the most abundant antigen on the erythrocyte-invading malaria parasite surface. The protein appears to be flexible and adopts an unusual fold with a large central cavity, which may be fundamental to its proposed role as a major hub for interaction with other MSPs and erythrocyte surface proteins. Previously identified RBC interaction sites could be mapped onto the flexible domain, as well as a strongly basic region in the central cavity. This and other resolved basic patches could also play a role in initial weak adhesion to erythrocytes via long-range electrostatic interactions, e.g., with highly negatively charged heparin-like polysaccharides, which are abundant on the erythrocyte surface. Notably, the coiled-coil fold of the p38 and p42 subunits hints at a role in membrane remodeling or cytoskeleton interactions, which has indeed been proposed previously for MSP-1 (24, 25). Dimerization of MSP-1, which involves coiled-coil regions, is hampered by the presence of spectrin, supporting this hypothesis. The structures presented here only form a small piece of the puzzle, given that MSP-1 forms a complex with other MSPs on the parasite surface. Nevertheless, the high-resolution detail afforded by our models provides an important framework to guide future work to fully elucidate the role of the native MSP-1 complex in Plasmodium infectivity and antibody-mediated protection from malaria.

MATERIALS AND METHODS

MSP-1 preparation

MSP-1 was produced recombinantly in Escherichia coli and purified under GMP-compatible conditions by Biomeva GmbH (Heidelberg, Germany) using the P. falciparum MSP-1D construct described by Kauth et al. (28). Briefly, the coding sequence of MSP-1D is largely based on the P. falciparum strain 3D7, but the N-terminal signal peptide and the C-terminal GPI anchor have been removed. The coding sequence was split into two halves (p83/30 and p38/42); each was expressed individually in E. coli W3110-Z1 and recovered as inclusion bodies. The recombinant MSP-1D was reconstituted as a heterodimer (hdMSP-1) in a process called “pulse renaturation” (39). The protein was purified using ion exchange and hydroxyapatite chromatography followed by dialysis into phosphate-buffered saline (PBS) and lyophilization. MSP-1 produced in this way was previously shown to assemble in the same manner as the native MSP-1 complex (28) (density observed for a disulfide bond between C813 and C845 in our EM maps and the proximity of C211 and C216 in the maps are also indicative of native folding). Before use, lyophilized hdMSP-1 was reconstituted by adding ddH2O, filtered through a 0.22-μm pore centrifugal filter, and subjected to SEC at 4°C on a Superdex 200 Increase 10/300 column (GE Healthcare) equilibrated with gel filtration buffer [25 mM Hepes and 125 mM NaCl (pH 7.4)]. SEC peak fractions were pooled and concentrated to ~0.7 to 2.5 mg/ml using an Amicon Ultra centrifugal concentrator [molecular weight cutoff (MWCO) 100,000; Merck]. For the production of fully SUB-1–cleaved MSP-1, hdMSP-1 was processed in vitro following the procedure published by Kousis et al. (40). Lyophilized hdMSP-1 was reconstituted by adding ddH2O and dialyzed into SUB-1 buffer [50 mM tris and 15 mM CaCl2 (pH 7.6)]. After dialysis, precipitated protein was removed by centrifugation. Recombinant P. falciparum SUB-1 (SUB-1), provided by M. J. Blackman, was added to hdMSP-1 in SUB-1 buffer, and the sample was incubated at 16°C overnight. Nonspecific cleavage and protein degradation were prevented by adding the protease inhibitors E64, leupeptin, and Pefabloc. Successful processing was verified.
by SDS-PAGE. The processed sample was filtered through a 0.22-μm pore centrifugal filter and again subjected to SEC at 4°C on a Superdex 200 Increase 10/300 column equilibrated with gel filtration buffer. SEC peak fractions were pooled and concentrated to ∼1.6 to 2.7 mg/ml using an Amicon Ultra centrifugal concentrator (MWCO 3000).

**EM sample preparation and cryo-EM data collection**

Concentrated hdMSP-1 and fully processed MSP-1 SEC peak fractions were diluted to ∼0.02 mg/ml and examined by uranyl formate negative staining using a Tecnai Spirit BioTWIN electron microscope at 120 kV to confirm sample homogeneity. For cryo-EM, Quantifoil R2/2 and R1.2/1.3 holey carbon grids (copper and gold 300 mesh, respectively) were washed in chloroform overnight before use. Grids were prepared in the presence and absence of 0.125% (w/v) CHAPS (3-[3-cholamidopropyl] dimethylammonio]-1-propanesulfonate) or subjected to iterative CTF refinement, 3D refinement, and Bayesian polishing until no further improvements in the EM maps were observed. Particle sets were then imported into cryoSPARC (42), subjected to 2D classification (which identified residual junk particles in only two cases) and final NU 3D refinement, resulting in six maps with nominal resolutions ranging from 3.1 to 3.6 Å (auto-tightened masked applied). The highest-resolution map was designated the “main” conformation. The alternative conformations (Alt conf) were numbered according to the position of the wing domain, from furthest (1) to closest (5) to the coiled-coil domain (the main conformation represents an intermediate position; see fig. S2A). To improve the resolution in the wing domain to aid model building, local refinement was performed in cryoSPARC (42) using the particle set corresponding to the main conformation. To this end, soft masks encompassing either H1-12 (wing domain) or H13-37 (“rest”) were generated in RELION 3.0 (43) using a map generated from the initially built atomic model in UCSF chimera (47) filtered to 12 Å. For local refinement of the wing domain, placing the fulcrum at the tip of the wing domain at the end of helices H10 and H11 gave the best results. For local refinement of the rest of the map, density corresponding to the wing domain was subtracted, and the fulcrum for refinement was placed at the center of mass of the mask. A graphical representation of the processing strategy used for the MSP-1 dataset is provided in fig. S2A.

For hdMSP-1, the four datasets were initially processed individually. Collected movies were curated either manually based on resolution estimates from CTFFIND 4.1 (45) and visual inspection (dataset 1) or using Warp (41). For dataset 1, 1,686 of 2,272 micrographs containing signal at 5 Å or better were retained. For dataset 2, 3,560 of 4,725 micrographs were retained, containing signal at 5.2 Å or better, with estimated defocus values ranging from −0.3 to −3 μm, and average motion in the first one-third of the frames of no more than 2 Å per frame. For dataset 3, 5,864 of 6,610 micrographs were retained, containing signal at 3.9 Å or better, with estimated defocus values ranging from −0.3 to −2.5 μm, and average motion in the first one-third of the frames of no more than 1 Å per frame. For dataset 4, 1,391 of 1,454 micrographs were retained, containing signal at 5 Å or better, with estimated defocus values ranging from −0.75 to −3 μm, and average motion in the first one-third of the frames of no more than 1 Å per frame. Subsequently, data were imported into RELION 3.0 (43). Motion correction and dose weighting were performed using MotionCor2 (44) using a 5 × 5 patch. CTF estimation was performed with CTFFIND 4.1 (45), after which 1.6 million particles picked by Warp (41) were extracted using a box size of 288 pixels at 1.073 Å per pixel. The particle set was imported into cryoSPARC (42) and cleaned by 2D classification. A majority of the 2D classes (4.4% of particles) corresponded to dimeric MSP-1 and was also excluded. The cleaned particle set was subjected to 3D heterogeneous refinement (classification) using a 3D reference generated ab initio using cryoSPARC (42) from ∼15,000 particles picked and preprocessed on-the-fly by Warp (41) during data collection. The resulting 3D classes (mostly defined by different conformations of the wing domain) were subjected to nonuniform (NU) refinement, followed by further cleaning by 2D and 3D classification. Particle star files were generated for the resulting particle sets using UCSF pyem (46), which were reimported into RELION 3.0 (43) and subjected to iterative CTF refinement, 3D refinement, and Bayesian polishing until no further improvements in the EM maps were observed. Particle sets were then imported into cryoSPARC (42), subjected to 2D classification (which identified residual junk particles in only two cases) and final NU 3D refinement, resulting in six maps with nominal resolutions ranging from 3.1 to 3.6 Å (auto-tightened masked applied). The highest-resolution map was designated the “main” conformation. The alternative conformations (Alt conf) were numbered according to the position of the wing domain, from furthest (1) to closest (5) to the coiled-coil domain (the main conformation represents an intermediate position; see fig. S2A). To improve the resolution in the wing domain to aid model building, local refinement was performed in cryoSPARC (42) using the particle set corresponding to the main conformation. To this end, soft masks encompassing either H1-12 (wing domain) or H13-37 (“rest”) were generated in RELION 3.0 (43) using a map generated from the initially built atomic model in UCSF chimera (47) filtered to 12 Å. For local refinement of the wing domain, placing the fulcrum at the tip of the wing domain at the end of helices H10 and H11 gave the best results. For local refinement of the rest of the map, density corresponding to the wing domain was subtracted, and the fulcrum for refinement was placed at the center of mass of the mask. A graphical representation of the processing strategy used for the MSP-1 dataset is provided in fig. S2A.

For hdMSP-1, the four datasets were initially processed individually. Collected movies were curated either manually based on resolution estimates from CTFFIND 4.1 (45) and visual inspection (dataset 1) or using Warp (41). For dataset 1, 1,686 of 2,272 micrographs containing signal at 5 Å or better were retained. For dataset 2, 3,560 of 4,725 micrographs were retained, containing signal at 5.2 Å or better, with estimated defocus values ranging from −0.3 to −3 μm, and average motion in the first one-third of the frames of no more than 2 Å per frame. For dataset 3, 5,864 of 6,610 micrographs were retained, containing signal at 3.9 Å or better, with estimated defocus values ranging from −0.3 to −2.5 μm, and average motion in the first one-third of the frames of no more than 1 Å per frame. For dataset 4, 1,391 of 1,454 micrographs were retained, containing signal at 5 Å or better, with estimated defocus values ranging from −0.75 to −3 μm, and average motion in the first one-third of the frames of no more than 1 Å per frame. Subsequently, data were imported into RELION 3.0 (43). Motion correction and dose weighting were performed using MotionCor2 (44) using a 5 × 5 patch. CTF estimation was performed with CTFFIND 4.1 (45). Particle picking was performed using cryOLO (48) for dataset 1 (yielding ∼219,000 particles) and Warp (41) for all other datasets (yielding ∼696,000, ∼1.2 million, and ∼304,000 particles for datasets 2 to 4, respectively). Particles were extracted using a box size of 288 pixels at 1.077 Å per pixel. Various processing strategies were tested for all datasets, and those resulting in the best maps were eventually followed; a graphical representation of the final processing strategy is provided in fig. S2B. Datasets were either first cleaned by 2D classification using ISAC from the SPHIRE cryo-EM software suite (49) or subjected directly to 3D classification in RELION 3.0 (43) using either unbinned data (box size, 288 pixels) or 4x binned data (box size, 72 pixels). Further rounds of 3D classification in RELION 3.0 (43)
were performed if necessary, followed by iterative CTF refinement, 3D refinement, and Bayesian polishing, until no further improvements to the maps were observed. Particles sets were then refined in 3D applying C2 symmetry (after having confirmed the symmetry in a 3D reconstruction without applying symmetry). No differences were observed in the final reconstructions of the four datasets at the obtained resolution beyond differences in the angular orientation distribution of the particles (fig. S2B), and thus, all particle sets were merged. Excluding any of the four datasets resulted in (slightly) worse final maps in subsequent processing steps. To better classify the conformational variability observed in the wing domain, the combined particle sets were subjected to symmetry expansion (50) and signal subtraction yielding a particle set corresponding to all monomers. The monomer mask used for signal subtraction was created in RELION 3.0 (43) using a monomer map generated from a dimer map in USCF Chimera (47) using the volume eraser tool. Further 3D classification without alignment into 10 classes using a monomer mask was performed. The best classes were refined individually in RELION 3.0 (43) yielding monomer maps with two distinct conformations of the wing domain. In a second approach, to get the best possible dimer maps, the best classes were also merged. Signal subtraction was reverted, and duplicates were removed. The resulting particle set was refined in 3D and subjected to 3D variability analysis in cryoSPARC (see movie S1) (S1). Two extreme conformations were extracted from the trajectory of the first mode of the variability analysis and used as references for multi-referenced 3D classification in cryoSPARC (42). This resulted in two maps with wing domain conformations similar to those obtained from RELION (requesting further classes supplying more references did not yield any additional good maps). Local resolution estimation of the maps was performed using cryoSPARC (42), except for the symmetry expansion–derived maps for which RELION 3.0 (43) was used. UCSF pyem (46) was used to generate Euler angle distribution plots for 3D reconstructions generated in cryoSPARC (42). All reconstructed maps had the wrong handedness and were flipped for atomic model building.

**Model building and analysis**

Buccaneer (52) was used to start de novo model building followed by manual building in Coot (53) guided by PSIPRED 4.0 (54) secondary structure prediction, using either RELION 3.0 (43) with automatic B-factor estimation to sharpen the maps or the auto-sharpened refined maps resulting from NU refinement in cryoSPARC (42). DeepEMhancer was also used to postprocess the map corresponding to the main conformation of MSP-1 (55). The locally refined maps [auto-sharpened in cryoSPARC (42) or sharpened using DeepEMhancer (55)] were also used in model building of the main conformation of MSP-1. For hdMSP-1, the C2 symmetric dimer maps showed higher local resolution for most of the protein, while the monomer maps obtained from the symmetry-expanded particle set showed a better-resolved wing domain (figs. S2 to S4) and were used to build that part. PHENIX (56) was used for model refinement, and the model was scored with Molprobity (57). The refined model was then used to locally sharpen the refined EM map using LocScale (58). The LocScale-sharpened EM map was subsequently used in combination with the other maps for further iterative refinement in Coot (53), ISOLDE (59), and PHENIX (56). The best resolved conformation for each sample was built initially, and the resulting atomic model was fit into the maps of the other conformation(s) using ISOLDE (59). Interhelical interactions of the flexible wing domain were restrained until the model had settled into the density of the alternative conformation, after which restraints were released and further iterative refinement in Coot (53), ISOLDE (59), and PHENIX (56) was performed. After confirming that the independently built MSP-1 and hdMSP-1 models did not differ substantially, the wing domain built for the main conformation of MSP-1 was fit into the hdMSP-1 maps and refined in the same manner. Map-model FSC (Fourier shell correlation) curves were generated using PHENIX (56).

The model was refined in the same manner against one of the independent half-maps and cross-validated against the both independent half-maps; the overlay of the resulting map-model FSCs indicated that no overrefinement took place. Further model quality statistics were determined using PHENIX (56) and can be found in table S1. Structural figures were prepared using UCSF Chimera (47), USCF ChimeraX (60), and Inkscape. An initial topology cartoon was generated by Pro-Origami (61) and modified in Inkscape. The ConSurf webserver (62) and USCF Chimera (47) were used to map the MSP-1 sequence conservation onto the derived structure using a sequence alignment of all *Plasmodium* MSP-1 genes available from PlasmoDB (63) generated by Clustal Omega 1.2.4 (64). The PDBsum webserver (65) and USCF Chimera (47) were used to analyze the dimerization interface. The FATCAT webserver (31) was used to identify known structural motifs in the determined MSP-1 structure.

**Cross-linking mass spectrometry**

Cross-linking combined with mass spectrometry (MS) analysis was used to validate the MSP-1 model. Lyophilized hdMSP-1 (150 μg) was reconstituted by adding 150 μl of ddH2O, followed by buffer exchange into PBS. A stock solution of disuccinimidyl sulfoxide (DSSO; Thermo Fisher Scientific) was freshly prepared according to the manufacturer’s instructions. After optimizing the cross-linker concentration by analyzing cross-linked hdMSP-1 by SDS-PAGE, hdMSP-1 was incubated with a 20-fold molar excess of DSSO for 60 min at room temperature before stopping the reaction with quenching buffer. Samples were digested in urea lysis buffer [8 M urea, 100 mM NaCl, and 50 mM triethylammonium bicarbonate (TEAB) (pH 8.5)] using a two-stage Lys-C and trypsin digestion protocol. Samples were reduced and alkylated using 10 mM dithiothreitol (1 hour, 27°C) and 30 mM iodacetamide (30 min, room temperature, in the dark). Lys-C was added at a 1:100 enzyme-to-protein molar ratio, and the sample was incubated at 37°C for 4 hours. Afterward, the sample was diluted 1:5 using 50 mM TEAB. Trypsin was added at a 1:50 enzyme-to-protein molar ratio, and samples were incubated overnight at 37°C. Digestion was stopped by acidification [0.6% (v/v) trifluoroacetic acid (TFA)] before desalting using StageTips (66). Dried peptides were resuspended in reconstitution buffer (2.5%, 1,1,1,3,3,3-hexafluoro-2-propanol and 0.1% TFA in water) before liquid chromatography (LC)–MS measurement, which was conducted using an ultra-performance liquid chromatography system (Ultimate 3000 UPLC) (Thermo Fisher Scientific) coupled to an Q-Exactive HF-X mass spectrometer (Thermo Fisher Scientific). During the LC separation, the peptides were first loaded onto a trapping cartridge (Acclaim PepMap300 C18, 5 μm, 300-Å-wide pore; Thermo Fisher Scientific) and washed for 3 min with 0.1% TFA in water. Analytical separation was performed using a nanoEase MZ Peptide analytical column (300 Å, 1.7 μm, 75 μm by 200 mm; Waters) and carried out for 90-min total analysis time. The 90-min LC method consisted of a multistep gradient going from 2 to 8% solvent B (80% acetonitrile).
and 20% water with 0.1% formic acid) in 5 min, 8 to 25% in 45 min, and 25 to 40% in 10 min followed by a washing and an equilibration step while solvent A was water and 0.1% formic acid. Eluting peptides were analyzed online by a coupled Q-Exactive-HF-X mass spectrometer running in data-dependent acquisition mode. Full scans were performed at 120K resolution on a mass range covering 375 to 1500 mass/charge ratio (m/z) with a maximum allowed injection time (maxIT) of 54 ms, followed by up to 8 tandem mass spectrometry (MS/MS) scans at 60K resolution with a maxIT of 100 ms for up to 1 \times 10^5 ions [automatic gain control (AGC) target]. Precursors were isolated with a window of 1.6 m/z and fragmented with a collision energy of 30 NCE (normalized collision energy). Unassigned, singly and doubly charged features were excluded from fragmentation, and dynamic exclusion was set to 10 s. Raw files were converted into mzML files using default settings in MzConvert (v3.0.21048) from the ProteoWizard toolkit (67). mzML files were then analyzed using Merox (v2.0.1.4.) (68) for identification of cross-linked peptides using a single protein database of the MSP-1D protein. Overall, default search parameters were used with the following changes: Cross-linker was set to DSSO, lysines were set as cross-linking sites, consecutive peptides were ignored as cross-linkers, and analysis mode was set to “quadratic mode” as recommended for small fasta files. A 1% false discovery rate cutoff was used including manual inspection of lower-scoring hits. The resulting table was exported as csv and used for 3D structural analysis using R and ChimeraX (60).

**Thermostability and MST characterization**

The thermal stability and aggregation behavior of (hd)MSP-1 were investigated using nanoscale differential scanning fluorimetry (nanoDSF) on a Prometheus NT.48 instrument (NanoTemper Technologies). Samples obtained after SEC were concentrated to \(-0.7\) to 2.7 mg/ml and loaded into nanoDSF-grade standard capillaries (NanoTemper Technologies). Samples were heated from 15 to 95°C at 1°C/min, and the intrinsic fluorescence at \(\lambda = 330\) nm after excitation at \(\lambda = 280\) nm and scattering were used to monitor receptor denaturation and aggregation, respectively. Relatively high protein concentrations were needed because of the lack of Trp residues in MSP-1. The first derivative of the unfolding curve was used to determine the transition midpoint using PR.ThermControl software (NanoTemper Technologies).

The affinity of spectrin for MSP-1 was assessed by MST. Spectrin (Sigma-Aldrich) was labeled with NT-647-NHS dye (NanoTemper Technologies). Samples were heated from 15 to 95°C at 1°C/min, and the intrinsic fluorescence at \(\lambda = 330\) nm after excitation at \(\lambda = 280\) nm and scattering were used to monitor receptor denaturation and aggregation, respectively. Relatively high protein concentrations were needed because of the lack of Trp residues in MSP-1. The first derivative of the unfolding curve was used to determine the transition midpoint using PR.ThermControl software (NanoTemper Technologies).

The affinity of spectrin for MSP-1 was assessed by MST. Spectrin (Sigma-Aldrich) was labeled with NT-647-NHS dye (NanoTemper Technologies). Samples were heated from 15 to 95°C at 1°C/min, and the intrinsic fluorescence at \(\lambda = 330\) nm after excitation at \(\lambda = 280\) nm and scattering were used to monitor receptor denaturation and aggregation, respectively. Relatively high protein concentrations were needed because of the lack of Trp residues in MSP-1. The first derivative of the unfolding curve was used to determine the transition midpoint using PR.ThermControl software (NanoTemper Technologies).
22. M. J. Boyle, J. S. Richards, P. R. Gilson, W. Chai, J. G. Beeson, Interactions with heparin-like...

23. T. J. Scott-Finnigan, A. A. Holder, Extensive proteolytic processing of the malaria parasite merozoite surface protein MSP1 that specifically binds to human erythrocytes...
61. A. Stivala, M. Wybrow, A. Wirth, J. C. Whisstock, P. J. Stuckey, Automatic generation of protein structure cartoons with pro-origami. Bioinformatics 27, 3315–3316 (2011).

62. H. Asklenazy, S. Abadi, E. Mertz, O. Chay, I. Mayrose, T. Pupko, N. Ben-Tal, ConSurf 2016: An improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res. 44, W344–W350 (2016).

64. F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, J. Rappsilber, M. Mann, Y. Ishihama, Protocol for micro-purification, enrichment, and final pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. 2, 129–134 (2007).

67. M. Blackman for providing SUB-1, B. Hessling for contributions to the XL-MS analysis, and M. Kudryashev for providing essential reagents. P. M. D. wrote the manuscript with input from the other authors. M. K. supervised the project and provided the project funding. Competing interests: H. B. is a shareholder of Sumaya GmbH & Co. KG. The other authors declare that they have no competing interests. Data and materials availability: The atomic coordinates of (hd)MSP-1 have been deposited at the Protein Data Bank (PDB) under accession codes 6ZBJ and 6ZBL for the monomeric hdMSP-1 conformations 1 and 2, and 6ZBC, 6ZBD, 6ZBF, 6ZBG, and 6ZBH for the dimeric hdMSP-1 conformations 1 and 2, and 6ZBC, 6ZBD, 6ZBF, 6Z BG, and 6ZBH for the monomeric main MSP-1 conformation and alternative conformations 1 to 5, respectively. The corresponding cryo-EM density maps have been deposited at the Electron Microscopy Data Bank (EMDB) under accession codes EMD-11156, EMD-11157, EMD-11150, EMD-11152, EMD-11151, EMD-11153, EMD-11154, and EMD-11155, respectively. The original movies and refined particle datasets have been deposited at the electron microscopy public image archive (EMPIAR; accession code EMPIAR-10437). The XL-MS data have been deposited to the proteomeXchange Consortium via the PRIDE (proteomics identifications) partner repository with the dataset identifier PXD024749.

71. L. H. Miller, T. Roberts, M. Shahabuddin, T. F. McCutchan, Analysis of sequence diversity in the Plasmodium falciparum merozoite surface protein-1 (MSP-1). Mol. Biochem. Parasitol. 59, 1–14 (1993).

Acknowledgments: We thank D. Mills and the EM staff at Max Planck Institute of Biophysics for expert technical assistance, J. F. C. Hernandez and O. Yildiz for computing support, M. Blackman for providing SUB-1, B. Hessling for contributions to the XL-MS analysis, and F. Frischknecht for critical reading of the manuscript. Funding: P. M. D., Y. Z., S.-Y. S. C., and M. K. were supported by the Sofja Kovalevskaja Award from the Alexander von Humboldt Foundation to M. K. Y. Z. was partially supported by an IMPRS international student scholarship. S.-Y. S. C. was partially supported by a GSSA student scholarship from the Taiwanese government. Author contributions: T. M. and H. B. provided the samples. T. M. and P. M. D. prepared the SUB-1–processed MSP-1. P. M. D. performed biophysical analyses, the optimization of EM grid preparation, EM data collection, processing, model building, and model analysis. Y. Z. contributed to the cryo-EM data collection and preliminary data analysis. S.-Y. S. C. contributed to the biophysical analyses. D. H. performed the XL-MS analysis. M. L. provided essential reagents. P. M. D. wrote the manuscript with input from the other authors. M. K. supervised the project and provided the project funding.