

A Scale-Up Compatible *Pichia pastoris* Platform for High-Yield Production of Recombinant *Plasmodium falciparum* Circumsporozoite Protein Variants

Neo Judith Lemme

Texas Children's Hospital Center for Vaccine Development

Rampa Victor Mosweu

Texas Children's Hospital Center for Vaccine Development

Ihonga Kuhepa

Texas Children's Hospital Center for Vaccine Development

Amandeep Kaur

Texas Children's Hospital Center for Vaccine Development

Shravani Deo

Texas Children's Hospital Center for Vaccine Development

Yi-Lin Chen

Texas Children's Hospital Center for Vaccine Development

Mei Du

Texas Children's Hospital Center for Vaccine Development

Rojelio Mejia

Baylor College of Medicine

Zhuyun Liu

Texas Children's Hospital Center for Vaccine Development

Wen-Hsiang Chen

Baylor College of Medicine

Jungsoon Lee

Texas Children's Hospital Center for Vaccine Development

Ulrich Strych

Baylor College of Medicine

Peter Hotez

Baylor College of Medicine

Maria Elena Bottazzi

Baylor College of Medicine

Nestor L. Uzcategui

nestor.uzcateguiaraujo@bcm.edu

Research Article

Keywords: circumsporozoite protein, malaria, process development, *Pichia pastoris*

Posted Date: October 15th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11220940/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

A Scale-Up Compatible *Pichia pastoris* Platform for High-Yield Production of Recombinant *Plasmodium falciparum* Circumsporozoite Protein Variants

Neo Judith Lemme^{1,2}, Rampa Victor Mosweu^{1,2}, Ihonga Kuhepa^{1,2}, Amandeep Kaur¹, Shravani Deo¹, Yi-Lin Chen¹, Mei Du¹, Rojelio Mejia³, Zhuyun Liu¹, Wen-Hsiang Chen^{1,3}, Jungsoon Lee¹, Ulrich Strych^{1,3}, Peter Hotez^{1,3,4,5}, Maria Elena Bottazzi^{1,3,4,5}, Nestor L. Uzcategui^{1,3,*}.

¹ Texas Children's Hospital Center for Vaccine Development, Houston, TX, USA

² Ministry of Health Botswana-Government of Botswana, Gaborone Botswana

³ Departments of Pediatrics and Molecular Virology and Microbiology, National School of Tropical Medicine Baylor College of Medicine, Houston, TX, USA

⁴ Baylor University, Department of Biology, Waco, TX, USA

⁵ Rice University, James A. Baker III Institute for Public Policy, Houston, TX, USA

*Corresponding author: Nestor L. Uzcategui (nestor.uzcateguiaraujo@bcm.edu)

Abstract:

With 282 million cases and 610,000 deaths annually, malaria represents a significant global health threat. The circumsporozoite protein (CSP), the predominant surface antigen of *Plasmodium falciparum* sporozoites, is the parent antigen employed to derive the major target antigens in the two licensed malaria vaccines, RTS,S/AS01 and R21/Matrix-M. Notably, the efficacy of these vaccines remains limited, underlining the need for improvement. This study describes a scalable, animal-free bioprocess to produce a tag-free recombinant CSP variant in the yeast *Pichia pastoris*. The protein, PfaCSP0002, corresponds to the antigen component used in the R21/Matrix-M vaccine, but is extended by Region I and its flanking amino acids. These additional epitopes have been suggested to improve protein stability and are recognized targets for protective monoclonal antibodies that confer passive immunity in animal models. The expression yield, purification, and protein quality of PfaCSP0002 were compared with those of a full-length CSP variant, PfaCSP0004. Both constructs were successfully expressed and purified, but PfaCSP0002 outperformed PfaCSP0004, reaching a fermentation yield of 399 mg/L. A final yield of 137 mg/L of 97–99% pure, monodisperse PfaCSP0002 protein was obtained after purification by cation exchange and size-exclusion chromatography. Both variants were specifically recognized by sera from malaria-infected individuals, with PfaCSP0002 showing higher reactivity than PfaCSP0004. This work lays the foundation for GMP-compliant production of CSP variants in *P. pastoris*, suitable for different purposes, including structural, functional, immunological research, or vaccine development.

Keywords: circumsporozoite protein, malaria, process development, *Pichia pastoris*

Introduction

Malaria remains a global health threat, with *Plasmodium falciparum* as the main species responsible for human deaths worldwide, particularly in children. Approximately 282 million malaria cases were reported globally in 2024, resulting in 610,000 deaths, with most cases occurring in sub-Saharan Africa (WHO, 2025).

Despite decades of effort, developing an effective, long-lasting protective malaria vaccine remains elusive due to the parasite's complex life cycle and ability to evade the immune system through mechanisms such as antigenic variability and hiding inside host cells (Sallam et al., 2025). An important focus in malaria vaccine research is the pre-erythrocytic stage of the parasite, especially the sporozoites' invasion of hepatocytes. The circumsporozoite protein, the predominant surface antigen of *P. falciparum* sporozoites (PfaCSP), is essential for this process and thus represents a key target for vaccine development. First cloned and expressed in the 1980s, this antigen has undergone further modification, derivation, and development to produce the malaria virus-like particle (VLP) vaccines, RTS,S and R21 (Collins et al., 2017; Ellis et al., 1983). Structurally, CSP is composed of three major domains: an N-terminal region containing T-cell epitopes and a conserved proteolytic cleavage site, a central repeat region characterized by immunodominant NANP and occasional NVDP motifs, and a C-terminal region that includes a thrombospondin type I repeat (TSP) motif, crucial for hepatocyte invasion and CD8⁺ and CD4⁺ cytotoxic cell epitopes that contribute to protective cellular immunity (Noe et al., 2014). The CSP-based RTS,S and R21 vaccines consist of the C-terminal domain of CSP, which includes half of the immunodominant central repeat region, fused to the surface protein of the hepatitis B virus. While these vaccines are recommended by the World Health Organization (WHO), their efficacy is still limited, ranging from 30-

50% for RTS,S to 63-76% for R21, following a 4-dose vaccination schedule (Collins et al., 2017). Interestingly, some N-terminal epitopes of CSP that are absent from the RTS,S and R21 vaccines have been shown to induce monoclonal antibodies that confer experimental protection against malaria (Atcheson et al., 2021). Therefore, newly engineered CSP variants remain key components for next-generation malaria vaccines, whether in VLPs, recombinant subunit combinations, or fusion chimeras with other antigens. The development of a robust and scalable production process for manufacturing these engineered CSP variants according to Good Manufacturing Practice (GMP) is important to support their progression through the stages of clinical development. Additionally, a high-yielding, optimized process for producing CSP variants will enable their use as reagents and support research into the immunogenic and functional roles of less-characterized protein regions. Such antigens could also be combined with other target malaria antigens or antigens from other pathogens causing co-endemic infections on the African continent.

This study demonstrates the feasibility of scalable production and purification of tag-free recombinant CSP variants in *P. pastoris* using animal-free reagents. Notably, a new variant of CSP, similar to that present in the R21 vaccine, but containing additional epitopes (PfaCSP0002), exhibited high expression and efficient purification. Importantly, the variant was specifically recognized by sera from malaria-positive individuals, with significantly elevated IgG levels compared to controls, even higher than those observed for the full-length protein, PfaCSP0004. This work established a flexible platform process accommodating structurally diverse CSP variants, providing a foundation for their use as reagents in biochemical and immunological studies and, ultimately, for potential vaccine development under GMP conditions.

Material and Methods

Design Rationale for PfaCSP Variants

For the design of the recombinant proteins, the CSP sequence from the *P. falciparum* NF54 strain (GenBank: KAF4328688.1) was used as the reference sequence. The DNA and amino acid sequences were analyzed using SnapGene (v8.2.2), ChimeraX (Meng et al., 2023; Pettersen et al., 2021), and AlphaFold 3 (Abramson et al., 2024) to support the design and structural evaluation of the resulting recombinant variants.

Cloning and Expression

The DNA encoding the CSP variants PfaCSP0001, PfaCSP0002 and PfaCSP0004 (**Fig. 1**) was codon-optimized for expression in *P. pastoris*, synthesized and cloned into the *PichiaPink* expression vector, pPink α -HC, using animal-free reagents. The three constructs provided by GenScript were linearized using the restriction enzyme AflII and electroporated into competent *P. pastoris* (*PichiaPink*TM strain #4; Thermo Fisher Scientific). The transformants were cultured on *Pichia* Adenine Dropout (PAD) agar plates and incubated at 30 °C for 4-5 days until colonies appeared. Ten colonies from each transformation were used for small-scale expression (10 ml), according to the manufacturer's instructions for the *Pichia* expression kit (Invitrogen). Briefly, they were first grown in Buffered Minimal Glycerol medium (BMGY) and then in Buffered Methanol-complex Medium (BMMY), with protein expression induced with 0.5% methanol daily at 30 °C for 3 days. After 72 hrs of induction, the cells were harvested. The supernatant was analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting using a monoclonal anti-circumsporozoite protein antibody (clone 2A10, Absolute Antibody). Clones with the highest protein expression were selected for glycerol seed stocks.

pH Screening analysis for PfaCSP protein expression

To identify the optimal pH for maximal PfaCSP protein expression in Buffered Minimal Glycerol (BMG) medium (similar to BMGY but without yeast extract and peptone; 13.4 g/L yeast nitrogen base, 1 g/L glycerol, 0.4 µg/mL biotin, 100 mM potassium phosphate pH 6.0), 1 mL of the seed stock was inoculated into a 200 mL BMG medium in a shake flask. The cells were expanded at 250 rpm overnight at 30 °C until the OD₆₀₀ reached between 1 and 10. For induction, cells were diluted to an OD₆₀₀ of 1 in 10 mL BMM medium (13.4 g/L yeast nitrogen base, 0.5% methanol, 0.4 µg/mL biotin, 100 mM potassium phosphate pH 6.0). Each 50 mL tube was filled with the required amount of seed culture, then centrifuged at 1500 × g for 5 minutes to pellet the cells. The supernatant was removed, and the cells were resuspended in 9 mL BMM. To adjust the pH, 1 mL of 1 M pH buffer (citrate buffer pH 3.9, 4.5, 5.4, 5.9 or phosphate buffer pH 6.3, 6.7, 7.0, 7.4) was added to each tube, while BMMY was included as a positive control. Methanol was added to a final concentration of 0.5% at 24 and 48 hours after inoculation. Samples (1 mL) were collected at 24, 48, and 72 hours after induction and saved for SDS-PAGE analysis. OD₆₀₀ and wet cell weight (WCW) were also measured at each time point.

Baseline Fermentation Process

One vial of the research cell bank was used to inoculate 0.5 L of BMG medium in a 2 L shake flask. The culture was incubated overnight at 30°C with shaking at 250 rpm until an OD₆₀₀ of 5–10. The volume of this culture necessary to inoculate the 1 L fermenter to a starting cell density of OD₆₀₀ = 0.5 was then calculated.

The fermentation was performed at the 1 L scale using a Sartorius BIOSTAT® Qplus bioreactor system. The fermentation process of PfaCSP0002 (Run 1 and 2) and PfaCSP0004 (Run 3 and 4) proteins started with an overnight glycerol batch phase for

cell expansion in 500 mL of Basal Salt Medium (BSM) (18.2 g/L potassium sulfate, 14.9 g/L magnesium sulfate heptahydrate, 4.13 g/L potassium hydroxide, 0.93 g/L calcium sulfate dehydrate, 26.7 mL/L of 85% phosphoric acid) with 30 g/L glycerol (Run 1 and 3) or 40 g/L glycerol (Run 2 and 4) at 30 °C with 30% dissolved oxygen until glycerol depletion. PTM1 trace salt and 200 mg/L of biotin were added during both the glycerol batch phase and the methanol induction phase. During the first hour of the methanol induction phase, the temperature was reduced from 30 to 25 °C, and the pH was adjusted from pH 5.0 to pH 4.8. For the induction phase, methanol was increased from 1 to 6 mL/L/hr (Runs 1 and 3) or from 1 to 8 mL/L/hr (Runs 2 and 4) over a 6-hour period. After this methanol adaptation phase, the methanol feed was maintained at 6 mL/L/h (Runs 1 and 3) or 8 mL/L/hr (Runs 2 and 4) until the end of fermentation. Total induction time was approximately 48 ± 2 hours.

Samples were taken daily, measuring biomass and OD₆₀₀. The fermentation culture was harvested by centrifugation at 12,227 x g for 30 minutes at 4 °C. The fermentation supernatant was filtered and stored at -80 °C for SDS-PAGE, densitometry analysis, and further purification.

Optimization of the Fermentation Process

To optimize the fermentation process, a second fermentation was performed to incorporate a glycerol fed-batch phase to increase biomass before methanol induction, and to test different induction pH values for PfaCSP0004. The fermentation process was similar to the baseline process except that upon glycerol depletion, 50% glycerol was added for a 5-hour fed-batch phase at a flow rate of 10 mL/L/hr (Runs 5 and 7) or 15 mL/L/hr (Runs 6 and 8). During the last hour of the fed-batch phase, the temperature was reduced from 30 to 25 °C, and the pH was adjusted from pH 5.0 to pH 4.8 (Runs 5 and

6) or pH 4.2 (Runs 7 and 8). For the induction phase, methanol was increased from 1 to 4 mL/L/hr (Runs 5 and 7) or 1 to 8 mL/L/hr (Runs 6 and 8) over a 6-hour period. After this methanol adaptation phase, the methanol feed was maintained at 4 mL/L/h (Runs 5 and 7) or 8 mL/L/hr (Runs 6 and 8) until the end of fermentation. Total induction time was approximately 68 ± 2 hours.

Development of the Purification Process

For the purification of the recombinant PfaCSP proteins, 500 mL of fermentation supernatant (FS) obtained from the fermentation process (Run 5 for PfaCSP0002 and Run 8 for PfaCSP0004) was thawed and buffer-exchanged with 20 mM citrate buffer at pH 4.0 by tangential flow filtration (TFF) using a Pellicon 3 cassette with a 10 kDa MWCO filter (Millipore). Cation exchange chromatography (CEX) was used to capture PfaCSP proteins using a CM Sepharose Fast Flow column (Cytiva) for PfaCSP0002 or SP Sepharose XL column (Cytiva) for PfaCSP0004. After loading, both proteins were eluted by a linear gradient from 0 to 1 M NaCl. Both proteins were further purified with size exclusion chromatography (SEC) using a Sephacryl S-200 column (Cytiva), pre-equilibrated in a 20 mM Tris-HCl, pH 7.5, 100 mM NaCl. PfaCSP0004 was concentrated using a centrifugal concentrator (Amicon Ultra-15 with 10 kDa MWCO) prior to SEC, while no concentration was needed for PfaCSP0002. Purity was analyzed by gel densitometry, and purified proteins were stored at -80°C until further usage.

Characterization of PfaCSP Proteins

Dynamic Light Scattering (DLS): DLS was used to determine the molecular weight, particle dispersion (%PD), and protein size (radius in nm) of the PfaCSP proteins. 1 mL of each sample, PfaCSP0002 (1.13 mg/ml) and PfaCSP0004 (0.61 mg/mL), was filtered through a $0.02\ \mu\text{m}$ filter (Whatman) to remove any interfering particles. 40 μL of each

sample was aliquoted into 10 wells of a 384-well plate (Corning, Cat# 3540). A Wyatt DynaPro plate reader II was used to analyze the protein particle size.

Size Exclusion High-Pressure Liquid Chromatography (SE-HPLC): The purity of the PfaCSP0002 and PfaCSP0004 proteins was evaluated using an SE-HPLC (Waters HPLC) system with an XBridge Premier column (Waters, Cat# 186009959). 100 µl of each sample, PfaCSP0002 (1.13 mg/mL) and PfaCSP0004 (0.61 mg/mL), were loaded onto the column and isocratically eluted with TBS (20 mM Tris-HCl, pH 7.5, 150 mM Sodium Chloride) at a flow rate of 0.5 mL/min. A gel filtration standard (Bio-Rad, Cat# 1511901) was used to evaluate the molecular weight of the purified proteins.

Western Blot: Western Blot was employed to confirm the PfaCSPs identity and evaluate impurities in the purified sample. To identify PfaCSP0002 and PfaCSP0004 proteins, anti-PfaCSP antiserum (1:3000 or 1:5000 dilution, Anti-Circumsporozoite protein [2A10], Rabbit IgG, Absolute antibody, Cat # Ab03231-23.0) and goat anti-Rabbit IgG secondary antibodies (1:5000 dilution, Invitrogen, Cat #31466) were used. To identify contaminating host proteins, an anti-host cell protein (HCP) antibody, Goat anti-*P. pastoris*, at a 1:2000 dilution in phosphate-buffered saline with 0.05% Tween 20 (PBST) was used.

ELISA-Based Immunogenicity Assessment

To evaluate the immunogenicity of the PfaCSP variants, de-identified human serum samples from donors infected with malaria in Uganda, obtained from Discovery Life Sciences, were analyzed, as previously described (Singampalli et al., 2022). The panel included 10 serum samples from individuals infected with *P. falciparum* (confirmed by HRP2 antigen detection) and 10 from non-traveling individuals from the USA, serving as negative controls. PfaCSP0002 and PfaCSP0004 recombinant proteins were assessed.

Enzyme-linked immunosorbent assay (ELISA) plates were coated with each antigen, and all serum samples were tested in duplicate. Optical density (OD) readings were obtained at 405 nm and 620 nm, and the net signal was calculated by subtracting background absorbance using the formula: $\Delta OD = (OD_{405} - \text{blank}) - (OD_{620} - \text{blank})$.

Results

Design Rationale for PfaCSP Variants

CSP-derived recombinant proteins are the main antigens in malaria vaccines, and our goal was to develop a process for producing improved CSP protein variants. We selected three variants for evaluation: PfaCSP0001, PfaCSP0002, and PfaCSP0004. *In silico* structure predictions for all variants suggest a structured C-terminal core that includes the TSP motif. This region contains one α -helix and two β -sheet secondary structures. The N-terminal region exhibits a high degree of intrinsic disorder and is therefore highly flexible, likely due to the presence of repetitive NANP motifs (**Fig. 1, S1, and S2**). PfaCSP0001 (**Fig. 1b**) corresponds to the CSP antigen component used in the R21 vaccine, whereas PfaCSP0004 (**Fig. 1c**) represents the full-length CSP lacking the signal peptide and GPI anchor. PfaCSP0002 was created by adding the N-terminal region I (RI) and its adjacent amino acids (LRKPKHKKLKQPADGNPDPNANPNVDP) to PfaCSP0001, introducing two new immunogenic epitopes, L9 and CIS43, that are absent in R21 and have been reported to elicit protective monoclonal antibodies against malaria in animal models (Atcheson et al., 2021; Jelínková et al., 2022, 2021; Kisalu et al., 2018; Nepal et al., 2025) (**Fig. S1c**). This modification may also contribute to improving stability by protecting against protease degradation (Zhang et al., 2008). Overall, we reasoned that fusing these

regions to PfaCSP0001 would improve stability, expression yield, and antigenic properties of the recombinant protein.

Small scale expression of PfaCSP0001, -2 and -4

PfaCSP0001, -2 and -4 were cloned in pPink α -HC. The vector was linearized and transformed into *P. pastoris*. Ten colonies per construct were induced for protein expression with 0.5% methanol. After induction, cells were centrifuged, and protein secretion into the supernatants was analyzed. SDS-PAGE and Western blots using an anti-CSP antibody showed strong expression for PfaCSP0002 and PfaCSP0004 at approximately 38 kDa and 60 kDa, respectively. PfaCSP0001 showed reduced expression with a molecular weight of 36 kDa (**Fig. 2**). The molecular weights of these PfaCSP recombinant protein variants were significantly higher than the expected theoretical values; 19.3 kDa, 21.5 kDa, and 38.2 kDa for PfaCSP0001, PfaCSP0002 and PfaCSP0004, respectively. In fact, this phenomenon was also reported for other recombinant PfaCSP variants, whether produced alone or as components of virus-like particles (Geens et al., 2024; Pendyala et al., 2023; Xia et al., 2023; Zhang et al., 2008). It is likely due to the repeated NANP motif, as these intrinsically disordered regions are known for reduced SDS binding, leading to an extended structure and atypical migration through the gel (Garg et al., 2026; Geens et al., 2024; Pendyala et al., 2023; Xia et al., 2023; Zhang et al., 2008). Clones with the highest recombinant protein expression from PfaCSP0002 and PfaCSP0004 variants were selected for seed stock preparation and stored at -80 °C for further studies. PfaCSP0001 was not selected due to its barely detectable expression levels in SDS-PAGE.

Upstream Process Development

pH screening process: To determine the optimal pH for recombinant protein production,

we evaluated the expression of PfaCSP0002 and PfaCSP0004 in *P. pastoris* across pH values ranging from 3.9 to 7.4. After three days of methanol induction, SDS-PAGE analysis showed visible bands corresponding to the previously determined molecular weights of both recombinant proteins under different pH conditions (**Fig. 3**). Based on the intensity of these bands, both pH 4.8 and pH 5.9 yielded the highest expression levels for PfaCSP0002. OD₆₀₀ measurements showed the highest cell density at pH 4.8. Therefore, pH 4.8 was chosen as the optimal induction pH for PfaCSP0002. For PfaCSP0004, optimal protein expression was observed between pH 4.8 and 5.9, with similar OD₆₀₀ values ranging from 10 to 12. Although the OD reading was highest at pH 5.4, this pH is too close to the protein's theoretical pI (pH 5.2). As protein solubility is typically lowest at its pI, the baseline fermentation for PfaCSP0002 was carried out at pH 4.8.

Fermentation: A total of eight fermentation runs were performed at the 1 L scale to develop the fermentation process (**Fig. 4 and Table 1**). The first four baseline fermentations (**Fig. 4A**) were set to determine the amount of glycerol in the glycerol batch phase and to test different methanol feeding rates. The fermentation runs with increased glycerol levels in the batch phase (40 vs. 30 g/L), and a higher methanol feeding rate (8 vs. 6 mL/L/hr) produced more biomass as measured by wet cell weight (**Fig. 5b**). After methanol induction, the biomass barely increased, suggesting growth had ceased (**Fig. 5b**). Methanol feeding was paused briefly, and there was no dissolved oxygen spike on the trend curve, indicating methanol accumulation in the cell culture, which was likely the root cause of cell deterioration. Therefore, all four baseline fermentation runs were stopped after two days of induction.

Fermentation samples were taken on day 0 (before methanol induction), day 1 and day 2

for expression analysis. Protein bands of expected sizes appeared after methanol induction on SDS-PAGE gels (as indicated by the red arrow in **Fig. 5a**). The protein yields for PfaCSP0002 increased from 7 ± 2 (Run 1) to 14 ± 5 mg/L (Run 2) measured by gel densitometry (**Table 1**). The protein yields of PfaCSP0004 were lower (4 ± 1 mg/L for Run 3 and 2 ± 1 mg/L for Run 4) and did not increase proportionally to the biomass, which may suggest protein solubility or stability problems. Overall, fermentation yields of all four baseline fermentations (Runs 1-4) were not robust, so further optimization to increase protein yields was needed.

The optimized fermentation process included a 5-hour glycerol-fed batch before methanol induction (**Fig. 4b**), which significantly increased the biomass before induction. A crucial and major improvement over the baseline fermentation was the sustained biomass accumulation upon methanol induction (**Fig. 6c**). The baseline fermentation for PfaCSP0004 was conducted at an induction pH of 4.8 and had a low protein yield. Therefore, the induction pH for PfaCSP0004 was changed to 4.2, one pH below the theoretical pI of 5.2, to help with protein solubility and stability. The results confirmed this hypothesis, and a better yield was achieved at pH 4.2 for PfaCSP0004. The expression of both PfaCSP proteins correlated with the growth of the biomass, showing a steady increase in fermentation samples collected at different time points after induction (**Fig. 6a and 6b**).

Three SDS-PAGE gels were analyzed to determine the average amount of protein produced per liter under the evaluated fermentation conditions. PfaCSP0002 grown with a glycerol feeding rate of 10 mL/L/hr and a methanol feeding rate of 4 mL/L/hr (Run 5) showed a yield of 279 ± 32 mg/L of FS, while PfaCSP0004 (Run 7) showed a yield of 119 ± 7 mg/L of FS. At a glycerol feeding rate of 15 mL/L/hr and a methanol feeding rate

of 8 mL/L/hr, both PfaCSP0002 (Run 6) and PfaCSP0004 (Run 8) exhibited better fermentation yields of 399 ± 38 mg/L of FS and 198 ± 11 mg/L of FS, respectively (**Table 1**). The total amount of PfaCSP0002 and PfaCSP0004 from a 1 L scale optimized fermentation was about 213 ± 20 mg (Run 6) and 95 ± 5 mg (Run 8), respectively.

Downstream Purification

Purification of PfaCSP0002 Protein: When a CEX column was tested to capture PfaCSP0002 (pI= 8.7), the protein was bound completely at pH 4.0. For purification, the FS was therefore buffer-exchanged into 20 mM citrate buffer, pH 4.0, using TFF and applied to a CM Sepharose Fast Flow column (CMFF). Bound PfaCSP0002 protein was eluted with a linear gradient from 0 - 1 M NaCl in 20 mM citrate buffer, pH 4.0. A single peak containing PfaCSP0002 protein was seen during elution (**Fig. 7b**), and SDS-PAGE analysis of elution fractions indicated two dominant bands at ~49 kDa, likely corresponding to PfaCSP0002 protein. An additional high-molecular-weight smear and several host cell proteins (HCPs) at ~40 kDa and 6-10 kDa were observed with the target protein. CEX elution peak fractions shown in the red dotted box were pooled.

When the PfaCSP0002 protein was further purified by SEC using a Sephacryl S-200 column, the chromatogram showed three peaks (**Fig. 7c**), with the second peak containing the PfaCSP0002 protein, as shown in a red dotted box. The second peak fraction also contained a ~65 kDa band on non-reduced SDS-PAGE, suggesting potential dimer formation (**Fig. 7c**). On the other hand, the first peak contained high molecular weight smears, ~40 kDa HCP, and some dimers (~65 kDa), while the third peak contained 6-10 kDa HCPs. Hence, the second peak fractions shown in the red dotted box were pooled for further analysis (**Fig. 7c**).

Purification of PfaCSP0004 Protein: The CEX column was also tested to capture the PfaCSP0004 protein. Although its isoelectric point is predicted to be 5.0, the protein can bind to the CEX column using its heparin-binding property located at the N-terminus (Ancsin and Kisilevsky, 2004; Zhao et al., 2016). Since heparin is a glycosaminoglycan with a high negative charge, PfaCSP0004 also bound to a CEX column at pH 4.0. For purification, the FS was buffer-exchanged into 20 mM citrate buffer, pH 4.0, using TFF and applied to an SP Sepharose XL column. Bound PfaCSP0004 protein was eluted with a linear gradient from 0 - 1 M NaCl in 20 mM citrate buffer, pH 4.0. The CEX chromatogram showed two merged peaks, and SDS-PAGE analysis of elution fractions indicated that the first peak contained PfaCSP0004 protein at ~60 kDa with HCPs in 28-50 kDa size and the second peak contained 6-10 kDa HCPs. Since two peaks were merged, all CEX elution fractions were pooled based on SDS-PAGE as shown in a red dotted box in **Fig. 8b**.

PfaCSP0004 protein was further purified by SEC using a Sephacryl S-200 column after concentrating the CEX pool. The SEC chromatogram showed two peaks (**Fig. 8c**), and SDS-PAGE analysis revealed that the first peak fractions contained PfaCSP0004 protein and its potential dimer form at ~100 kDa, as shown in a red dotted box in **Fig. 8c**, while other fractions, including the second peak, also contained PfaCSP0004 protein but co-eluted with HCPs (6-38 kDa) and, hence, were excluded from pooling to make a final purified PfaCSP0004 protein pool with higher purity.

Characterization of PfaCSP Proteins

Analysis of Purified PfaCSP Proteins Using SDS-PAGE: After purification, 5 µg of the purified PfaCSP proteins were analyzed by reduced and non-reduced SDS-PAGE, and purity was measured by gel densitometry (**Fig. 9a and b**). Purified PfaCSP0002 protein

appeared as a single band with 99.0% purity in a reduced condition and as two bands in a non-reduced condition, suggesting the presence of a cysteine-mediated dimer. The purity of monomeric PfaCSP0002 was 90.1% in a non-reduced condition (**Fig. 9a**). Purified PfaCSP0004 appeared as a single band with 98.6% purity under reduced conditions and as two bands under non-reduced condition, also suggesting cysteine-mediated dimer formation, although less pronounced than for PfaCSP0002. The purity of monomeric PfaCSP0004 was 97.9% in a non-reduced condition (**Fig. 9b**).

During the production process, the yield and recovery of the PfaCSP proteins were monitored (**Fig. 9c**). The yield of PfaCSP0002 was 279 mg protein/L FS, and the yield after purification was 137 mg protein/L FS, resulting in ~50% recovery after purification. The yield of PfaCSP0004 was 198 mg protein/L FS, and the yield after purification was 29 mg protein/L FS, resulting in ~15% recovery after purification. Overall, the lower yield from PfaCSP0004 purification was due to the need to eliminate HCPs at 6-38 kDa.

Size and Purity Analysis by HPLC and DLS: HPLC equipped with a Waters XBridge Premier column was used to analyze the size and purity of the purified PfaCSP0002 and PfaCSP0004 proteins (**Fig. 10a**). The results showed that both PfaCSP0002 and PfaCSP0004 proteins were of high purity, i.e., 97% and 96%, respectively. Interestingly, when comparing the eluted peaks of these two purified proteins with the Bio-Rad Gel Filtration Standards, their molecular weights fell between 44 and 158 kDa, which were higher than the theoretical values, 21.5 and 38.2 kDa for PfaCSP0002 and PfaCSP0004, respectively.

To eliminate any potential interaction between protein molecules and the column matrix, DLS analysis was employed, as it directly examined the Brownian motion of particles in solution to populate the diffusion coefficient and the size of the protein. On DLS (**Fig.**

10b), the radius of the PfaCSP0002 size was determined as 3.26 nm, and that of PfaCSP0004 as 3.69 nm, corresponding to 53.3 kDa and 71.3 kDa for PfaCSP0002 and PfaCSP0004, respectively. These results revealed that the measured molecular weight remained higher than the predicted values. Nevertheless, the low polydispersity values (%PD <20%) measured by DLS for both purified PfaCSP0002 (7.0%) and PfaCSP0004 (1.2%) proteins indicate that the samples were monodispersed, which continued to suggest both proteins were well purified.

Seroreactivity of Human Sera Against Recombinant PfaCSP Variants

To evaluate the reactivity of the purified malaria vaccine antigen candidates, we assessed the seroreactivity of human sera from malaria-infected individuals by ELISA, using the recombinant proteins PfaCSP0002 and PfaCSP0004 as antigens. Sera from malaria-positive individuals showed significantly higher reactivity against PfaCSP0002 compared to controls (**Fig. 11a**). A similar result was observed for PfaCSP0004; malaria-positive samples showed greater reactivity than controls, although to a lesser extent than for PfaCSP0002 (**Fig. 11b**). When directly comparing samples from malaria-positive individuals, PfaCSP0002 was more strongly recognized by sera than PfaCSP0004 (**Fig. 11c**). Overall, these results suggest that both proteins are suitable for further vaccine development.

Discussion

This study demonstrates an animal-reagent-free production platform for recombinant CSP variants in *P. pastoris* with potential for industrial-scale manufacturing. Yeast expression systems, particularly *P. pastoris*, represent a scalable and cost-effective platform that has been routinely and successfully used for the production of vaccines and

other recombinant proteins in compliance with GMP regulatory requirements. The *Pichia* expression system has been licensed to more than one hundred biotech, pharmaceutical, and food companies. Additionally, this system has the advantage of being considered GRAS (Generally Recognized as Safe) by the FDA (Moraes et al., 2024). In this study, we used *in silico* analysis to design several *Plasmodium* CSP variants and established a *P. pastoris* research seed stock using exclusively animal-free reagents. Furthermore, we developed and optimized a scalable fermentation and purification process, ensuring both the seed stock and the production workflow are suitable for future technology transfer and GMP manufacturing. The process delivered the highly pure recombinant proteins, PfaCSP0002 and PfaCSP0004, suitable for structural, functional, and immunological studies. Downstream, such antigens could be combined with our portfolio of non-malarial parasite antigens for the purposes of developing innovative combination vaccines. Increasingly, policymakers are seeking mechanisms to build NextGen vaccines targeting multiple pathogens (Fatima and Hong, 2025).

One of the key goals of this work was to improve the stability and expression of CSP-derived recombinant proteins. Improving stability and reducing extracellular proteolysis increases recoverable yield, reduces formation of truncated CSP products, and simplifies downstream purification. Maintaining protein integrity will also protect antigen quality and reduce manufacturing variability and costs. To address these challenges, we designed a modified CSP construct incorporating an N-terminal peptide intended to improve recombinant protein stability (Zhang et al., 2008). We generated the PfaCSP0002 variant by adding an N-terminal stabilizing peptide to PfaCSP0001. PfaCSP0001 is the CSP variant in the established malaria vaccines (R21 and RTS,S), which consists of the C-terminus of the protein, as well as 19 NANP repeat motifs as the N-terminus. In small-

scale expression, PfaCSP0002 was expressed robustly, whereas PfaCSP0001 was barely expressed. Another truncated variant containing the N-terminal stabilizing peptide, PfaCSP0003, was generated to further evaluate its effect on protein recovery (**Fig. S3a and b**). Similar to PfaCSP0002, PfaCSP0003 showed high-level expression, further supporting a role for the peptide in enhancing protein yield (**Fig. S3c**). Thus, the peptide positively influenced protein yield, consistent with previous reports (Zhang et al., 2008). In addition, we used a protease-deficient *Pichia* strain (PichiaPink™ strain #4), which reduces proteolytic degradation of labile proteins. Together, these results suggest that the strong expression of PfaCSP0002 and PfaCSP0003 may result not only from the shielding of protease-sensitive NANP motifs but also from enhanced protein folding or secretory efficiency via mechanisms that have yet to be elucidated.

Interestingly, this region of the N-terminal stabilizing peptide, which corresponds to the junction between the N-terminal and central repeat regions of CSP, has also recently been associated with novel epitopes that generated protective monoclonal antibodies (L9 and CIS43) against malaria in animal models (Atcheson et al., 2021; Jelínková et al., 2021; Kisalu et al., 2018). This region alone, coupled with different VLPs, showed highly protective efficacy against malaria (Jelínková et al., 2022; Nepal et al., 2025). In a first step, we therefore evaluated the immunogenic potential of PfaCSP0002 and PfaCSP0004 by ELISA using human sera of infected and non-infected individuals. Both PfaCSP0002 and PfaCSP0004 were significantly better recognized by sera from malaria-infected individuals than by sera from controls, indicating that the recombinant variants preserve epitopes targeted during natural infection. Interestingly, PfaCSP0002 showed stronger reactivity than the full-length recombinant protein, PfaCSP0004, when testing the same malaria-positive samples against both antigens. These results suggest that the engineered

N-terminal extension could increase the exposure of relevant epitopes, whereas PfaCSP0004, with its longer and more disordered structure, could hide important epitopes present in the N-terminal stabilizing peptide, originally present in this recombinant protein. Although the sample size is small and functional antibody activity was not assessed, these findings support the antigenic authenticity of the recombinant proteins. Particularly for PfaCSP0002, which emerges as an attractive antigen warranting further characterization of the epitope specificity and protective potential against malaria, either as a stand-alone recombinant antigen or incorporated into VLP, similar to the R21 malaria vaccine.

A fermentation process was optimized to enable reproducible expression of multiple CSP variants by minimal adjustment of a common bioprocess framework. As proof of concept, we worked with two different variants, PfaCSP0002 and PfaCSP0004 (**Fig. 1, S1, and S2**). Baseline fermentations revealed plateaued cell growth upon methanol induction due to poor methanol adaptation and methanol toxicity. These inefficiencies were overcome by increasing the initial glycerol concentration from 30 to 40 g/L and incorporating a 5-hour glycerol fed-batch phase to enhance biomass accumulation prior to methanol induction. In addition, increasing the methanol feeding rate from 6 to 8 mL/L/hr provided both induction and energy for continued cellular metabolism during the production phase, further boosting biomass accumulation and overall recombinant protein yield. Under these optimized conditions, PfaCSP0002 and PfaCSP0004 achieved maximum yields of 399 ± 38 mg/L and 198 ± 11 mg/L, respectively. These values are compatible with industrial-scale vaccine antigen production and with the generation of high-quality reagents for research. In fact, PfaCSP0002 fermentation yield is highly comparable to those obtained during our SARS-CoV-2 vaccine development process (≥ 400 mg/L in a 1-

5 L bioreactor) (Hotez et al., 2023). This yield enabled large-scale manufacturing of the COVID-19 vaccines Corbevax (Biological E, India) and IndoVac (Bio Farma, Indonesia), together delivered around 100 million doses (Hotez et al., 2023).

We also demonstrate that a common purification workflow is effective for both CSP recombinant variants, which differ markedly in amino acid sequence, length, and predicted isoelectric point. Both variants could be captured by cation exchange chromatography at pH 4.0 and subsequently polished by size exclusion chromatography. The resulting preparations of both proteins achieved a monodisperse population (%PD < 20% by DLS) with > 95% purity by SDS-PAGE and SE-HPLC. The ability to process multiple CSP variants through a single streamlined purification workflow reduces the development burden for future variants. Again, the yield of PfaCSP0002, after purification, is highly comparable to our SARS-CoV-2 vaccine development process (137 mg/L for PfaCSP0002 vs 140 mg/L for SARS-CoV-2 recombinant vaccine) (Hotez et al., 2023). Taken together, these findings demonstrate the feasibility of the rapid implementation and production of new CSP variants with potential for industrial scalability.

Of note, the molecular weights of the recombinant PfaCSP proteins were significantly higher than their expected theoretical values, as calculated using different methods. This phenomenon has been previously reported in the literature for other recombinant PfaCSP variants, whether produced independently or as components of virus-like particles (Geens et al., 2024; Pendyala et al., 2023; Xia et al., 2023; Zhang et al., 2008). Although these differences could be due to post-transcriptional modifications, they are more likely to be associated with intrinsically disordered regions, such as the NANP sequences. Such regions typically exhibit weak SDS binding, resulting in a more elongated conformation

and, consequently, slower electrophoretic mobility (Garg et al., 2026; Geens et al., 2024; Pendyala et al., 2023; Xia et al., 2023; Zhang et al., 2008).

Although PfCSP and its engineered variants have been extensively investigated since the 1980s and serve as the main antigen in commercialized malaria vaccines (RTS,S and R21), published, scalable bioprocesses for these antigens are surprisingly limited. A comparative benchmarking of the process development workflows for PfaCSP0002 and PfaCSP0004 established in this study against representative PfCSP production systems reported over the past 25 years shows that our platform is highly competitive (**Table 2**) (Kolodny et al., 2001; Noe et al., 2014; Plassmeyer et al., 2009; Singh et al., 2020). The only evaluated parameter that surpassed our process was the recombinant protein yield achieved by the *Pseudomonas fluorescens* system reported by Noe *et al.* (2014); however, that technology is patented. Of note, we have previously shown that expression yields comparable to those achieved here are sufficient for the scalable, cost-effective production of recombinant COVID-19 vaccines (Hotez et al., 2023). Across all other evaluated quality and bioprocessing metrics, PfaCSP0002 and PfaCSP0004 consistently outperform alternative published platforms (Table 2). Importantly, our entire upstream and downstream process is executed strictly using animal-derivative-free reagents. Eliminating animal-derived materials mitigates the risk of transmitting infections, most notably transmissible spongiform encephalopathies (TSEs) and bovine spongiform encephalopathy (BSE). This practice aligns directly with regulatory guidelines and recommendations set forth by both the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA).

This study facilitates the systematic exploration of improved antigen designs that are both GMP-aligned and compatible with scalable production of structurally diverse CSP

variants, such as mutant and truncated forms, multi-antigen formulations, and VLP-incorporated constructs. Beyond vaccine development, the availability of highly pure, large amounts of CSP variants is valuable for a broad range of immunological and research applications. Although these proteins are not intended for use in primary diagnostics, they can serve as standardized antigens for ELISA or multiplex assays in sero-epidemiological studies conducted in the field. In addition, they are useful research tools, e.g., serving as substrates for monoclonal antibody isolation and epitope mapping, and for dissecting structure–function relationships, among other applications. Because the expression, fermentation, and purification protocols were implemented using well-defined process conditions, they are readily adaptable to producing additional CSP-based reagents in laboratories with limited expertise in bioprocess development.

In summary, we created an animal-free glycerol seed stock and established a scalable bioprocess for the production of tag-free CSP variants and therefore is of potential use in future anti-malarial or combination vaccines, and beyond. Additionally, we demonstrate its applicability using two different constructs. A newly engineered variant, PfaCSP0002, showed superior process development performance and greater recognition by sera from malaria-positive individuals compared with the full-length recombinant protein, PfaCSP0004. While further work is required to define the protective efficacy, optimal formulation, and potential for incorporation into VLPs of this antigen, the process reported here provides a practical foundation for CSP-focused vaccine development and immunological studies.

Acknowledgements:

We are grateful for the support of the academic training program from Texas Children's Hospital Global Health Network. We thank Prof. Peter B Lillehoj for providing the human serum samples from malaria-infected donors used in this study.

References

- Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., Bodenstein, S.W., Evans, D.A., Hung, C.C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., Beattie, C., Bertolli, O., Bridgland, A., Cherepanov, A., Congreve, M., Cowen-Rivers, A.I., Cowie, A., Figurnov, M., Fuchs, F.B., Gladman, H., Jain, R., Khan, Y.A., Low, C.M.R., Perlin, K., Potapenko, A., Savy, P., Singh, S., Stecula, A., Thillaisundaram, A., Tong, C., Yakneen, S., Zhong, E.D., Zielinski, M., Židek, A., Bapst, V., Kohli, P., Jaderberg, M., Hassabis, D., Jumper, J.M., 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 630, 493–500. <https://doi.org/10.1038/s41586-024-07487-w>
- Ancsin, J.B., Kisilevsky, R., 2004. A binding site for highly sulfated heparan sulfate is identified in the N terminus of the circumsporozoite protein: Significance for malarial sporozoite attachment to hepatocytes. *Journal of Biological Chemistry* 279, 21824–21832. <https://doi.org/10.1074/jbc.M401979200>
- Atcheson, E., Hill, A.V.S., Reyes-Sandoval, A., 2021. A VLP for validation of the *Plasmodium falciparum* circumsporozoite protein junctional epitope for vaccine development. *NPJ Vaccines* 6. <https://doi.org/10.1038/s41541-021-00302-x>

- Collins, K.A., Snaith, R., Cottingham, M.G., Gilbert, S.C., Hill, A.V.S., 2017. Enhancing protective immunity to malaria with a highly immunogenic virus-like particle vaccine. *Sci. Rep.* 7. <https://doi.org/10.1038/srep46621>
- Ellis, J., Ozaki, L., Gwadz, R., Cochrane, A., Nussenzweig, V., Nussenzweig, R., Godson, N., 1983. Cloning and expression in *E. coli* of the malarial sporozoite surface antigen gene from *Plasmodium knowlesi*. *Nature* 302, 536–538.
- Fatima, M., Hong, K.J., 2025. Innovations, Challenges, and Future Prospects for Combination Vaccines Against Human Infections. *Vaccines* (Basel). <https://doi.org/10.3390/vaccines13040335>
- Garg, A., Gielnik, M.B., Kjaergaard, M., 2026. Sequence determinants of the hypomobility of intrinsically disordered proteins in SDS-PAGE. <https://doi.org/10.64898/2026.03.24.714011>
- Geens, R., Stanisich, J., Beyens, O., D'Hondt, S., Thiberge, J.M., Ryckebosch, A., De Groot, A., Magez, S., Vertommen, D., Amino, R., De Winter, H., Volkov, A.N., Tompa, P., Sterckx, Y.G.J., 2024. Biophysical characterization of the *Plasmodium falciparum* circumsporozoite protein's N-terminal domain. *Protein Science* 33. <https://doi.org/10.1002/pro.4852>
- Hotez, P.J., Adhikari, R., Chen, W.H., Chen, Y.L., Gillespie, P., Islam, N.Y., Keegan, B., Tyagi Kundu, R., Lee, J., Liu, Z., Kimata, J.T., Oezguen, N., Pollet, J., Poveda, C., Razavi, K., Ronca, S.E., Strych, U., Thimmiraju, S.R., Versteeg, L., Villar-Mondragon, M.J., Wei, J., Zhan, B., Bottazzi, M.E., 2023. From concept to delivery: a yeast-expressed recombinant protein-based COVID-19 vaccine

technology suitable for global access. *Expert Rev. Vaccines* 22, 495–500.

<https://doi.org/10.1080/14760584.2023.2217917>

Jelínková, L., Flores-Garcia, Y., Shapiro, S., Roberts, B.T., Petrovsky, N., Zavala, F., Chackerian, B., 2022. A vaccine targeting the L9 epitope of the malaria circumsporozoite protein confers protection from blood-stage infection in a mouse challenge model. *NPJ Vaccines* 7. <https://doi.org/10.1038/s41541-022-00457-1>

Jelínková, L., Jhun, H., Eaton, A., Petrovsky, N., Zavala, F., Chackerian, B., 2021. An epitope-based malaria vaccine targeting the junctional region of circumsporozoite protein. *NPJ Vaccines* 6. <https://doi.org/10.1038/s41541-020-00274-4>

Kisalu, N.K., Idris, A.H., Weidle, C., Flores-Garcia, Y., Flynn, B.J., Sack, B.K., Murphy, S., Schön, A., Freire, E., Francica, J.R., Miller, A.B., Gregory, J., March, S., Liao, H.X., Haynes, B.F., Wiehe, K., Trama, A.M., Saunders, K.O., Gladden, M.A., Monroe, A., Bonsignori, M., Kanekiyo, M., Wheatley, A.K., McDermott, A.B., Farney, S.K., Chuang, G.Y., Zhang, B., Kc, N., Chakravarty, S., Kwong, P.D., Sinnis, P., Bhatia, S.N., Kappe, S.H.I., Sim, B.K.L., Hoffman, S.L., Zavala, F., Pancera, M., Seder, R.A., 2018. A human monoclonal antibody prevents malaria infection by targeting a new site of vulnerability on the parasite. *Nat. Med.* 24, 408–416. <https://doi.org/10.1038/nm.4512>

Kolodny, N., Kitov, S., Vassell, M., Miller, V., Ware, L., Fegeding, K., De La Vega, P., Sacci Jr, J., Lanar, D., 2001. Two-step chromatographic purification of recombinant *Plasmodium falciparum* circumsporozoite protein from *Escherichia coli*, *Journal of Chromatography B*.

- Meng, E.C., Goddard, T.D., Pettersen, E.F., Couch, G.S., Pearson, Z.J., Morris, J.H., Ferrin, T.E., 2023. UCSF ChimeraX: Tools for structure building and analysis. *Protein Science* 32. <https://doi.org/10.1002/pro.4792>
- Moraes, L.M.P. de, Marques, H.F., Reis, V.C.B., Coelho, C.M., Leitão, M. de C., Galdino, A.S., Porto de Souza, T.P., Piva, L.C., Perez, A.L.A., Trichez, D., de Almeida, J.R.M., De Marco, J.L., Torres, F.A.G., 2024. Applications of the Methylophilic Yeast *Komagataella phaffii* in the Context of Modern Biotechnology. *Journal of Fungi*. <https://doi.org/10.3390/jof10060411>
- Nepal, Y., Francian, A., Flores-Garcia, Y., Roberts, B.T., David, S.A., Zavala, F., Chackerian, B., 2025. Virus-like particle vaccines targeting a key epitope in circumsporozoite protein provide sterilizing immunity against malaria. *NPJ Vaccines* 10. <https://doi.org/10.1038/s41541-025-01241-7>
- Noe, A.R., Espinosa, D., Li, X., Coelho-Dos-Reis, J.G.A., Funakoshi, R., Giardina, S., Jin, H., Retallack, D.M., Haverstock, R., Allen, J.R., Vedvick, T.S., Fox, C.B., Reed, S.G., Ayala, R., Roberts, B., Winram, S.B., Sacci, J., Tsuji, M., Zavala, F., Gutierrez, G.M., 2014. A full-length *Plasmodium falciparum* recombinant circumsporozoite protein expressed by *Pseudomonas fluorescens* platform as a Malaria vaccine candidate. *PLoS One* 9. <https://doi.org/10.1371/journal.pone.0107764>
- Pendyala, G., Calvo-Calle, J.M., Moreno, A., Kane, R.S., 2023. A multivalent *Plasmodium falciparum* circumsporozoite protein-based nanoparticle malaria vaccine elicits a robust and durable antibody response against the junctional

epitope and the major repeats. *Bioeng. Transl. Med.* 8.

<https://doi.org/10.1002/btm2.10514>

Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I.,

Morris, J.H., Ferrin, T.E., 2021. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science* 30, 70–82.

<https://doi.org/10.1002/pro.3943>

Plassmeyer, M.L., Reiter, K., Shimp, R.L., Kotova, S., Smith, P.D., Hurt, D.E., House,

B., Zou, X., Zhang, Y., Hickman, M., Uchime, O., Herrera, R., Nguyen, V., Glen, J., Lebowitz, J., Jin, A.J., Miller, L.H., MacDonald, N.J., Wu, Y., Narum, D.L.,

2009. Structure of the *Plasmodium falciparum* circumsporozoite protein, a leading malaria vaccine candidate. *Journal of Biological Chemistry* 284, 26951–26963.

<https://doi.org/10.1074/jbc.M109.013706>

Sallam, Malik, Al-Khatib, A.O., Al-Mahzoum, K.S., Abdelaziz, D.H., Sallam,

Mohammed, 2025. Current Developments in Malaria Vaccination: A Concise Review on Implementation, Challenges, and Future Directions. *Clin. Pharmacol.*

<https://doi.org/10.2147/CPAA.S513282>

Singampalli, K.L., Li, J., Lillehoj, P.B., 2022. Rapid magneto-enzyme-linked

immunosorbent assay for ultrasensitive protein detection. *Anal. Chim. Acta* 1225.

<https://doi.org/10.1016/j.aca.2022.340246>

Singh, S.K., Plieskatt, J., Chourasia, B.K., Singh, V., Bolscher, J.M., Dechering, K.J.,

Adu, B., López-Méndez, B., Kaviraj, S., Locke, E., Richter King, C., Theisen, M.,

2020. The *Plasmodium falciparum* circumsporozoite protein produced in

Lactococcus lactis is pure and stable. *Journal of Biological Chemistry* 295, 403–414. <https://doi.org/10.1074/jbc.RA119.011268>

WHO, 2025. World Malaria Report 2025. World Health Organization.

Xia, M., Huang, P., Vago, F., Jiang, W., Tan, M., 2023. Pseudovirus Nanoparticles Displaying Plasmodium Circumsporozoite Proteins Elicited High Titers of Sporozoite-Binding Antibody. *Vaccines (Basel)*. 11. <https://doi.org/10.3390/vaccines11111650>

Zhang, Q., Ding, F., Xue, X., Xu, X., Pan, W., 2008. Changing the N-terminal sequence protects recombinant Plasmodium falciparum circumsporozoite protein from degradation in Pichia pastoris. *Appl. Microbiol. Biotechnol.* 78, 139–145. <https://doi.org/10.1007/s00253-007-1294-0>

Zhao, J., Bhanot, P., Hu, J., Wang, Q., 2016. A comprehensive analysis of Plasmodium circumsporozoite protein binding to hepatocytes. *PLoS One* 11. <https://doi.org/10.1371/journal.pone.0161607>

Statements & Declarations

Funding

Neo Judith Lemme, Rampa Victor Mosweu, and Ihonga Kuhepa received financial support from the Ministry of Health, Government of Botswana, and the Botswana Baylor Children's Clinical Centre of Excellence, Gaborone, Botswana, for their training visit to the Center for Vaccine Development at Texas Children's Hospital, Houston, Texas, USA. In addition, this work was partially supported by ExxonMobil.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Author Contributions

Study conception and design: MEB, PH, US, and NLU. *In silico* studies, molecular biology, and protein expression: NJL, RVM, IK, AK, and NLU. Fermentation, purification, and characterization of recombinant proteins: NJL, RVM, IK, SD, ZL, JL, YLC, MD, and WHC. ELISA assays: RM. Data analysis: ZL, JL, RM, WHC, and NLU. Manuscript (first draft): NJL, RVM, IK, and NLU. Manuscript revisions: WHC, JL, ZL, NLU, and US. Final manuscript preparation: NLU, US, PH, and MEB. All authors reviewed, edited, and approved the final manuscript.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files. Materials generated during the current study are available upon reasonable request.

Ethics approval

This study utilized commercially obtained, de-identified human serum samples. No identifiable personal information was available to the investigators. Ethical approval was not required.

Consent to participate

Not applicable. This study utilized commercially obtained, de-identified human serum samples.

Consent to publish

Not applicable.

Fig. legends

Fig. 1 Novel *P. falciparum* CSP vaccine antigen candidates based on the NF54 strain. **A)** The native PfaCSP protein (amino acids 1-397), which contains highly conserved RI motifs at the N-terminus, a central NANP repeat region, and a TSP/Region II (RII) near the C-terminus. **B)** The PfaCSP0001 variant (amino acids 193-373), corresponding to the sequence used in the R21 vaccine, which contains the TSP/RII region and half of the central repeat region (19 NANP repeats), lacking both the N-terminal domain and the C-terminal GPI anchor. **C)** PfaCSP0002 corresponds to the CSP fragment of PfaCSP0001 (amino acids 193-373), plus the N-terminal stabilizing peptide (amino acids 86-112), which includes the RI region and its adjacent amino acids. **D)** PfaCSP0004 (amino acids 18-373) consists of the full-length CSP sequence, minus the N-terminal signal peptide and the C-terminal GPI anchor. Pink and gray colors correspond to CD4⁺ and CD8⁺ epitopes, respectively.

Fig. 2 Small-scale expression analysis of PfaCSPs variants visualized using SDS-PAGE and Western blot. A, B and C represent PfaCSP0001, PfaCSP0002 and PfaCSP0004 samples, stained with Coomassie Blue. D, E and F represent PfaCSP0001, PfaCSP0002 and PfaCSP0004 samples, respectively, in Western blots. Lanes 1-10 represent samples of different clones expressing the recombinant protein. M: Marker. For the Western Blots, an anti-CSP primary antibody (clone 2A10, Absolute Antibody) was used.

Fig. 3 Optimization of pH conditions to maximize protein expression. Coomassie Blue-stained SDS-PAGE (left), OD600 measurements (right) for PfaCSP0002 (A) and PfaCSP0004 (B) pH screening analysis. Red triangles indicate the band of the recombinant protein. pH 4.8 was chosen for the fermentation of both proteins (red

rectangle). Day 1, Day 2 and Day 3 are samples taken after 24 hr, 48 hr and 72 hr of methanol induction

Fig. 4 Representation of the baseline (A) and optimized (B) fermentation process for recombinant proteins PfaCSP0002 and PfaCSP0004. The figure represents the schematic outline of the upstream workflow from inoculation of shake flasks to methanol induction, harvest, and clarification steps used for producing PfaCSP0002 and PfaCSP0004 in *P. pastoris*. The optimized steps or conditions are highlighted in blue.

Fig. 5 Expression (A) and biomass analysis (B) of baseline fermentation runs 1-4 for PfaCSP0002 and PfaCSP0004. From day 0 (before methanol induction) until day 3. Samples from pH screening in 10 mL BMM, pH 4.8 at day 3 were included as a positive control.

Fig. 6. Expression and yield analysis of four optimized fermentation runs 5-8 for PfaCSP0002 (A) and PfaCSP0004 (B) from day zero (before glycerol fed batch and before methanol induction) until day three. Wet cell weight for both PfaCSP proteins (C).

Fig. 7 Purification of PfaCSP0002. Purification scheme (A), chromatograms, and fraction analysis by reduced SDS-PAGE from the capture step (B) and by non-reduced SDS-PAGE from the polish step (C) are shown. The PfaCSP0002 elution peak and elution fraction analysis are shown in red dotted boxes. Black arrows indicate the PfaCSP0002 protein band.

Fig. 8 Purification of PfaCSP0004. Purification scheme (A) and chromatograms and fraction analysis by reduced SDS-PAGE from the capture step (B) and by non-reduced SDS-PAGE from the polish step (C) are shown. PfaCSP0004 elution peak and elution fraction analysis are shown in red dotted boxes. Black arrows indicate the PfaCSP0004 protein bands.

Fig. 9 SDS-PAGE analysis for the purified PfaCSP0002 (A) and PfaCSP0004 (B) proteins in reduced (R) and non-reduced (NR) conditions. (C) The purification yield, recovery, and purity were summarized.

Fig. 10 SE-HPLC chromatograms (A) and DLS (B) showing MW and purity of PfaCSP0002 and PfaCSP0004 proteins.

Fig. 11 Seroreactivity of human sera against recombinant PfaCSP variants measured by ELISA. Human serum samples from malaria-positive and malaria-negative (non-travelers) individuals were tested against the recombinant proteins PfaCSP0002 and PfaCSP0004 using an indirect ELISA.

Fig. 1

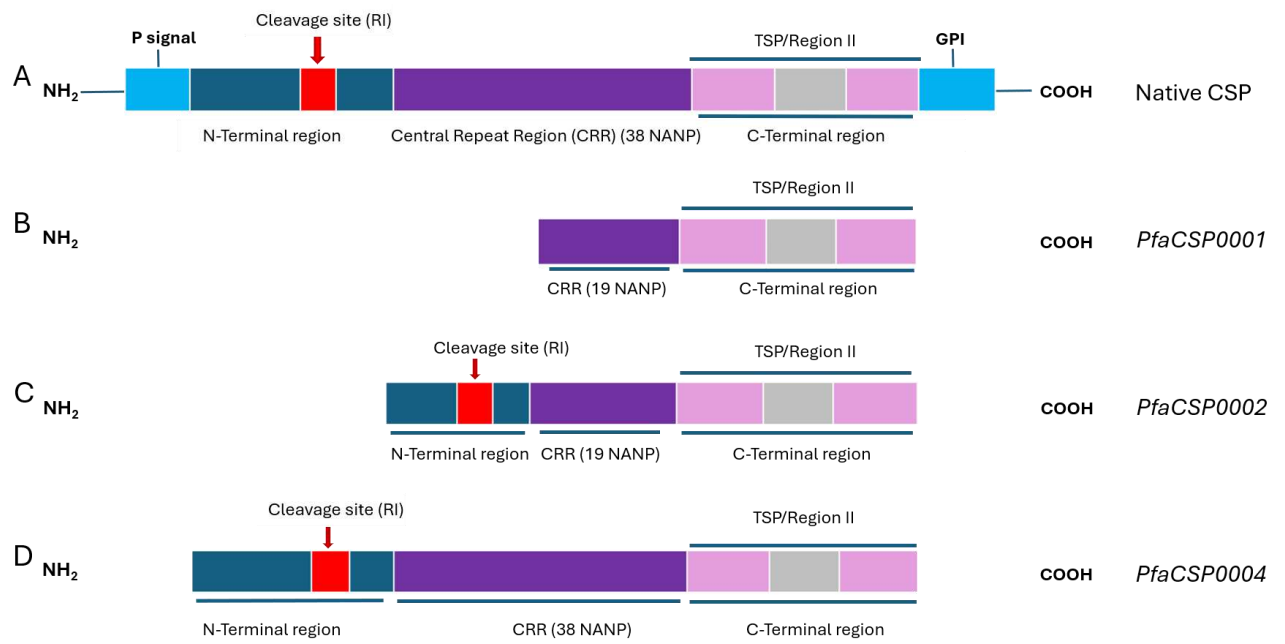


Fig. 2

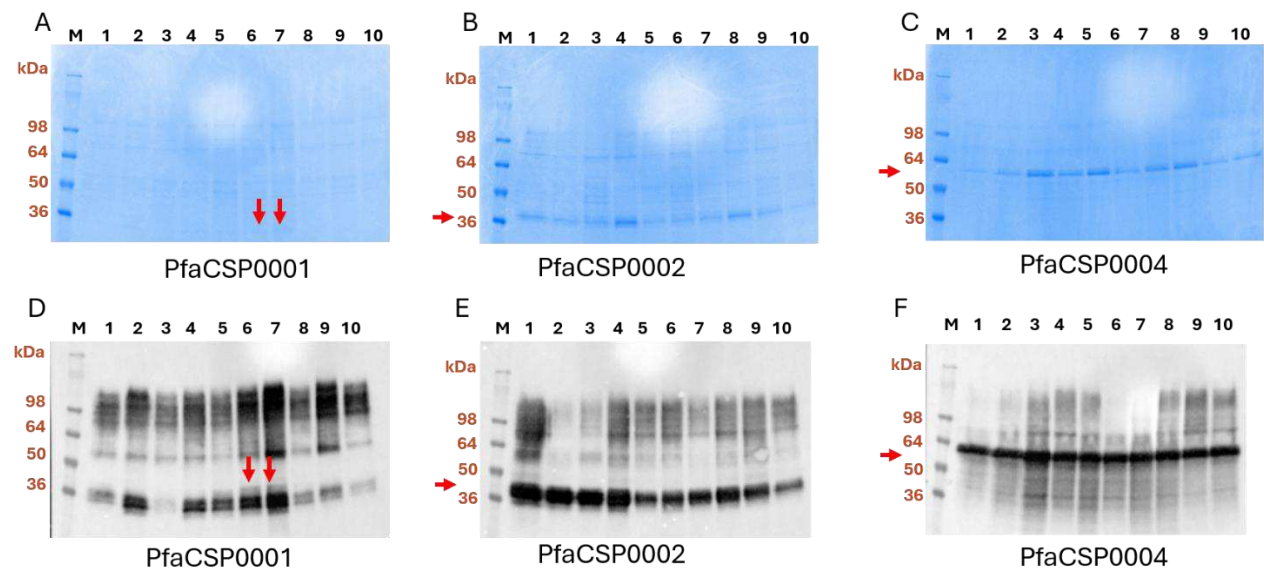


Fig. 3

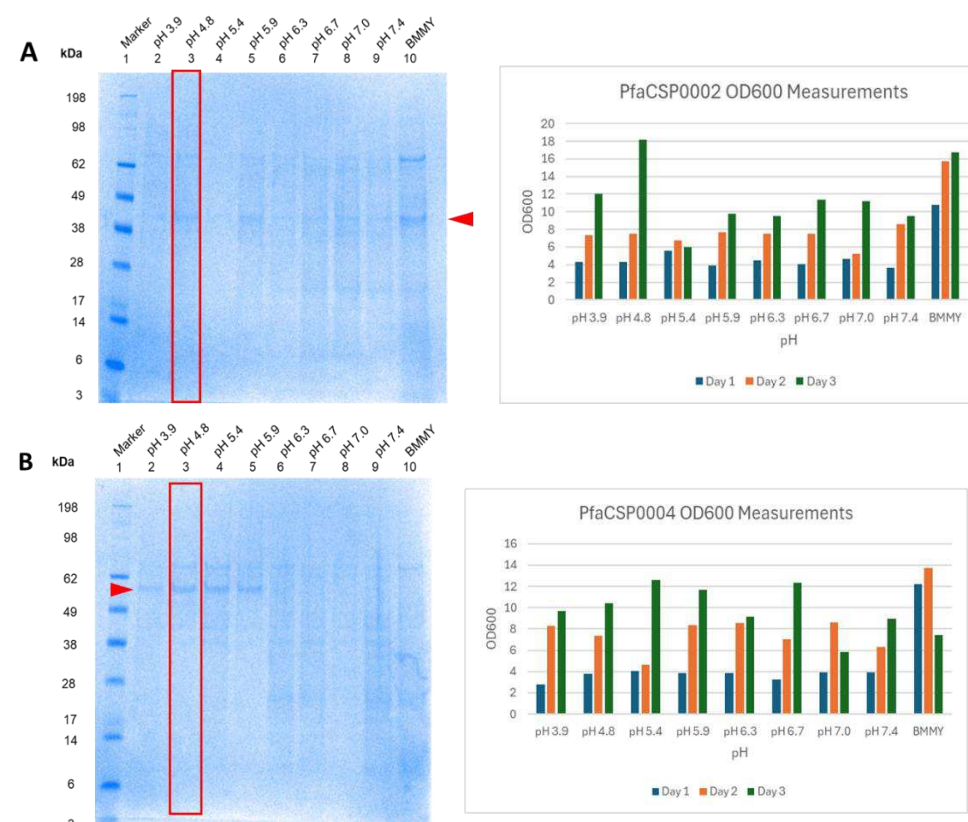


Fig. 4

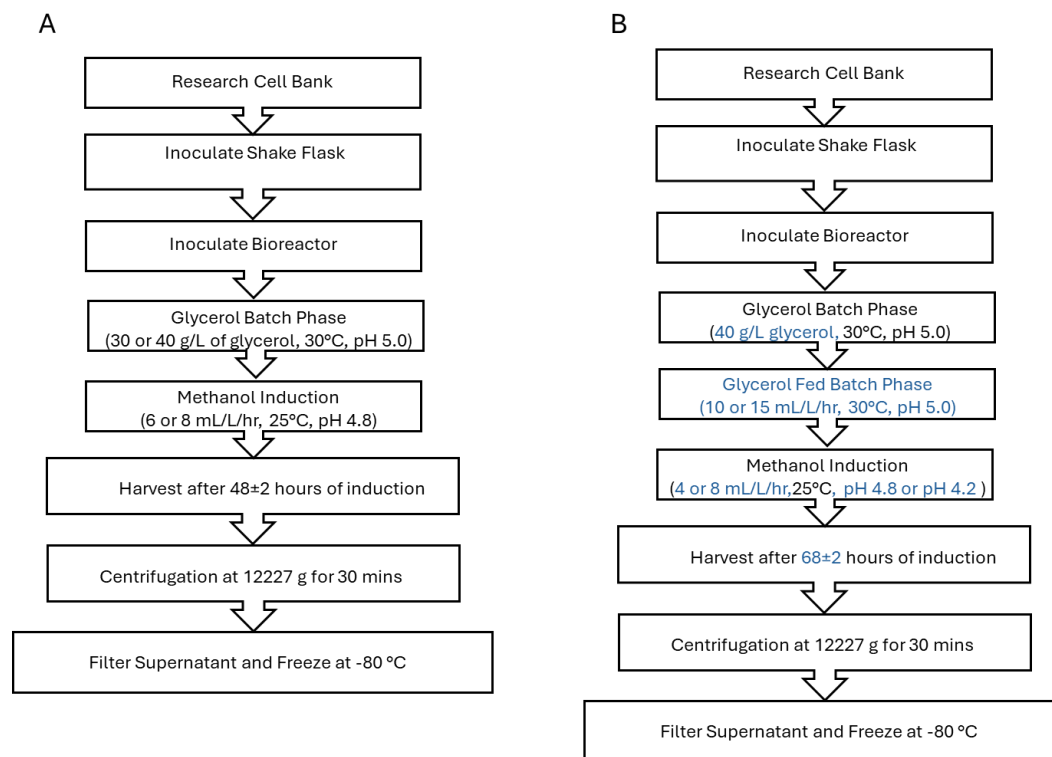


Fig. 5

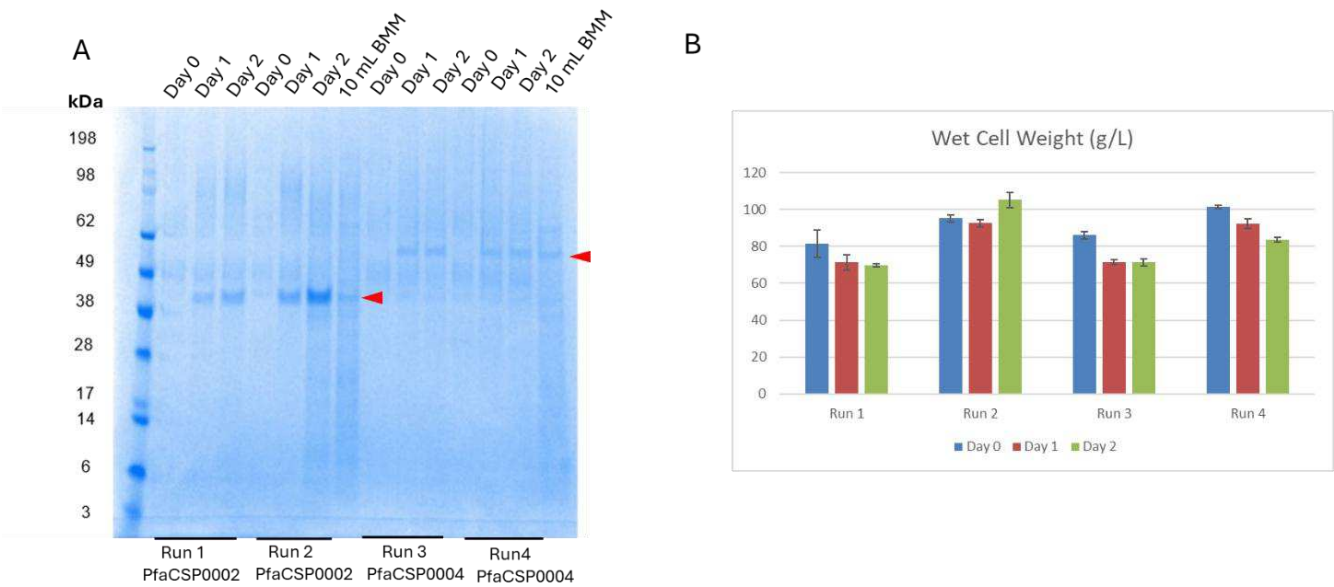


Fig. 6

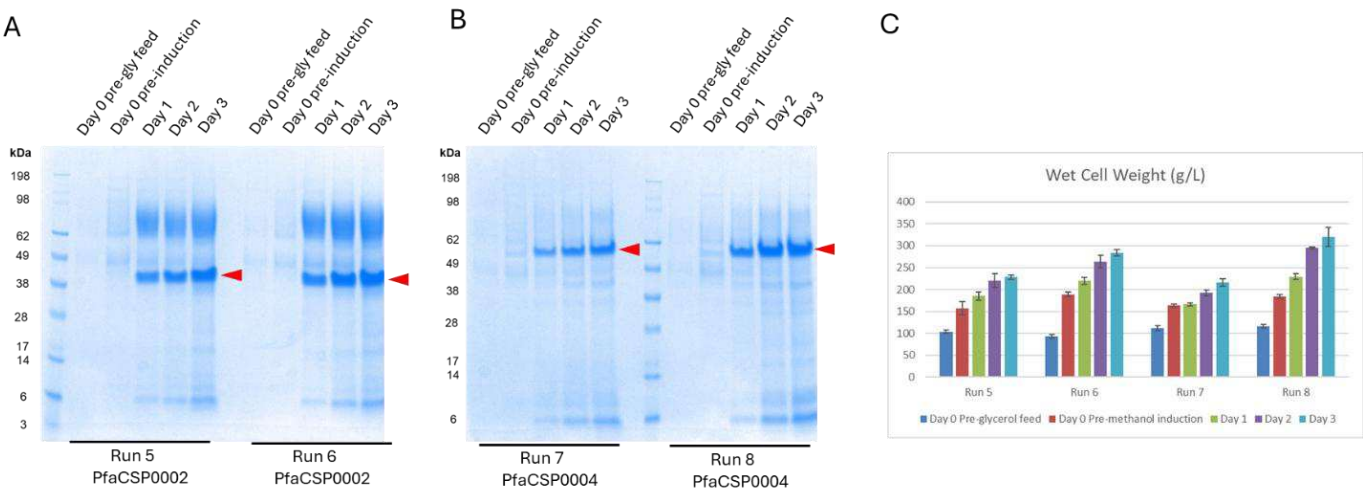


Fig. 7

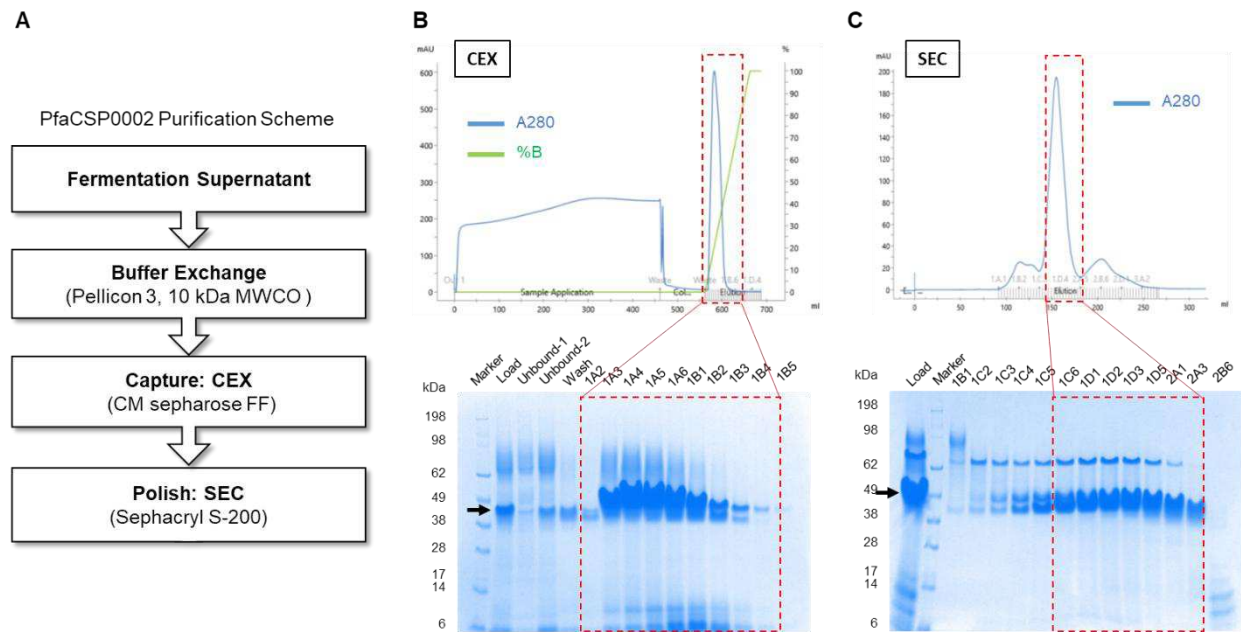


Fig. 8

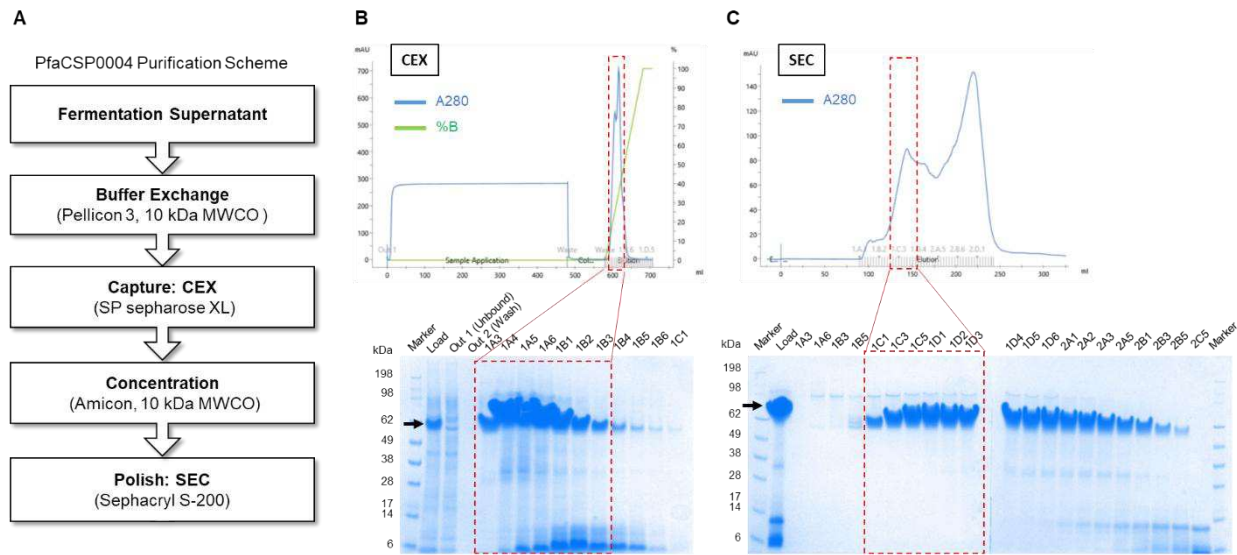


Fig. 9

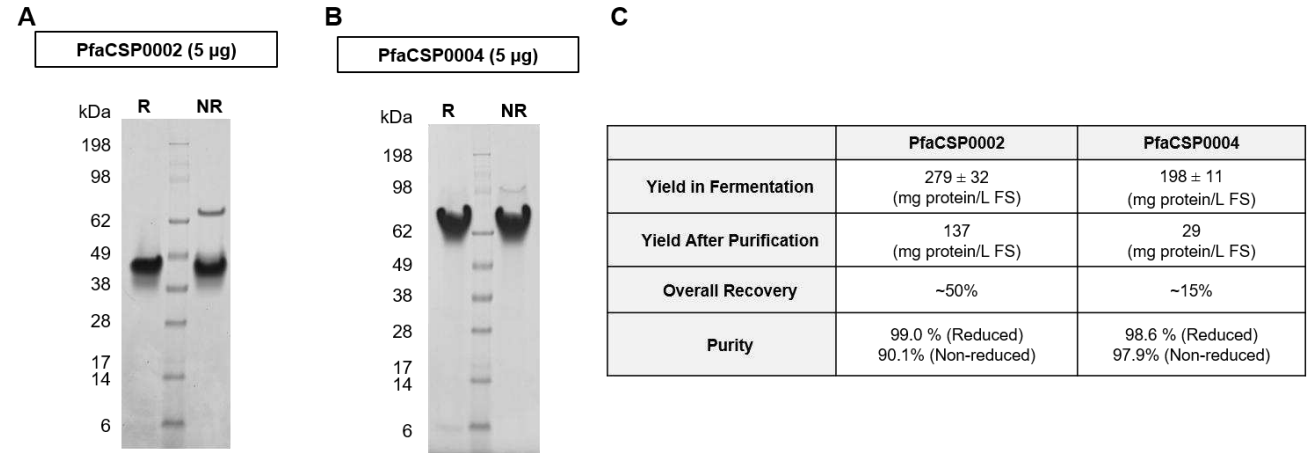


Fig. 10

A

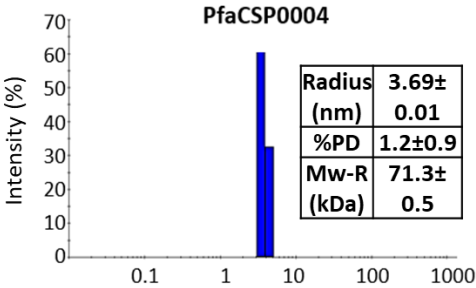
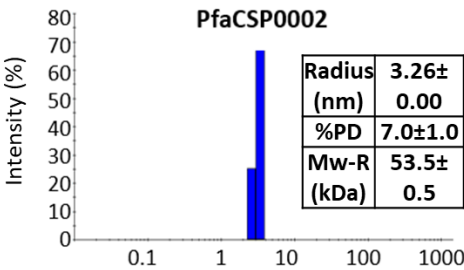
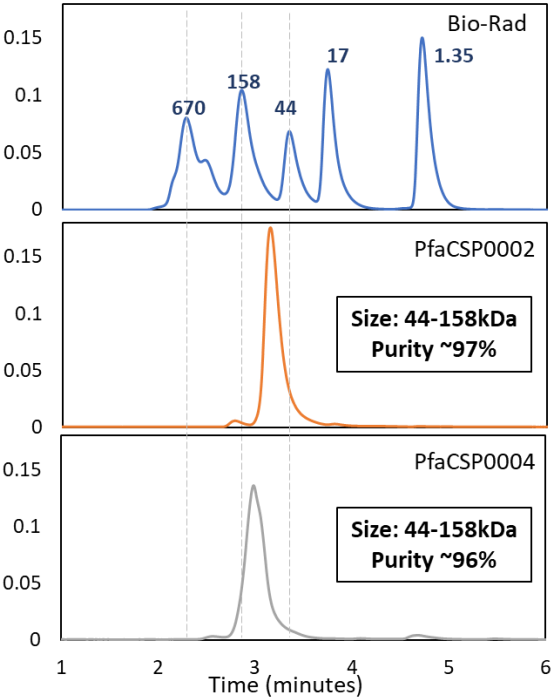


Fig. 11

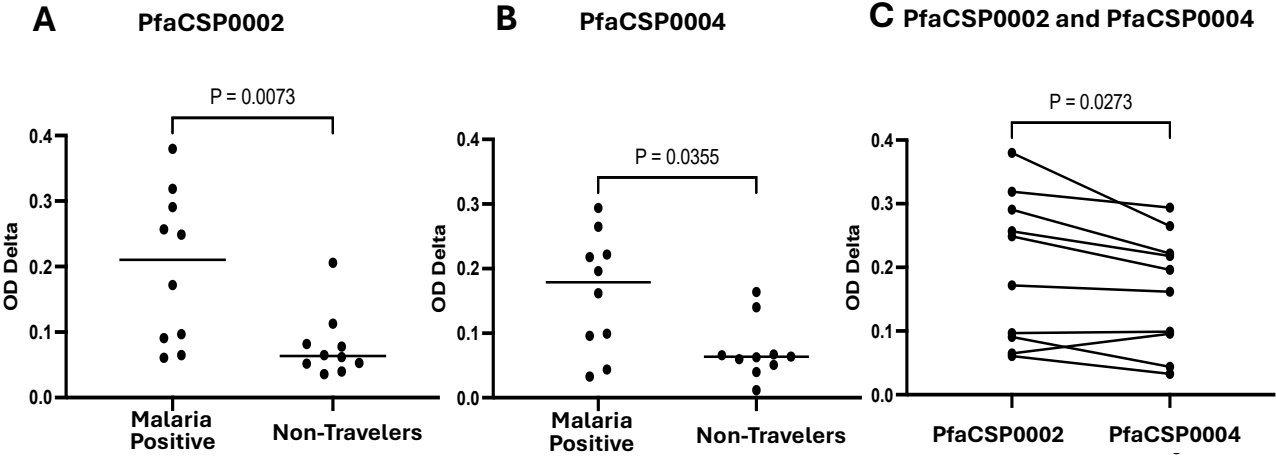


Table.

Table 1 Fermentation yields for PfaCSP0002 and PfaCSP0004

Run ID	Molecule	Glycerol in batch (g/L)	Glycerol feeding rate (mL/L/hr)	Methanol feeding rate (mL/L/hr)	Induction pH	WCW (g/L) at harvest	Yield (mg per L of FS)	Volume of FS (mL) From 1L run	Protein Yield (mg) from 1 L run
Run 1	PfaCSP0002	30	n/a	6	4.8	69.9 ± 0.9	7 ± 2	534	4 ± 1
Run 2	PfaCSP0002	40	n/a	8	4.8	105.2 ± 4.1	14 ± 5	485	7 ± 2
Run 3	PfaCSP0004	30	n/a	6	4.8	71.4 ± 2.0	4 ± 1	535	2 ± 1
Run 4	PfaCSP0004	40	n/a	8	4.8	83.6 ± 1.1	2 ± 1	534	1 ± 1
Run 5	PfaCSP0002	40	10	4	4.8	228.7 ± 5.1	279 ± 32	480	134 ± 15
Run 6	PfaCSP0002	40	15	8	4.8	283.4 ± 7.2	399 ± 38	534	213 ± 20
Run 7	PfaCSP0004	40	10	4	4.2	216.2 ± 9.2	119 ± 7	491	58 ± 3
Run 8	PfaCSP0004	40	15	8	4.2	320.2 ± 21.3	198 ± 11	480	95 ± 5

Table 2 Bioprocess Overview of PfaCSP0002 and PfaCSP0004: A Comparison of Published Recombinant PfCSP Platforms.

Recombinant protein (RP)	PfaCSP0002	PfaCSP0004	PfCSP4/38	CSP-1	EcCSP	PpCSP	rPfCS (aa 19-405)
PfCSP Antigen Design	R21 CSP antigen (aa 193–373) + peptide (86-112)	Full-length CSP (aa 18-373)	Full-length CSP (aa 26-383)	Full-length CSP (aa 21-382)	Full-length CSP (aa 19-411)	Near full-length CSP (aa 85-411)	Full-length CSP (aa 19-405)
purification tag	No	No	Yes (His tag)	No	Yes (His tag)	No	Yes (His tag)
Expression Host	<i>P. pastoris</i>	<i>P. pastoris</i>	<i>L. lactis</i>	<i>P. fluorescens</i>	<i>E. coli</i>	<i>P. pastoris</i>	<i>E. coli</i>
Expression Format	Soluble secreted RP	Soluble secreted RP	Soluble secreted RP	Soluble RP	Partially Soluble RP	Soluble secreted RP	Inclusion-body
Fermentation / expression Yield	Yes / 399mg/L	Yes / 198mg/L	Yes/ NR	Yes / 4g/L	Yes / NR	Yes / NR	Yes / NR
Yield after purification	137 mg/L	29 mg/L	25 mg/L	NR	NR	NR	20 mg/10g of wet cell
Purity	0.99	0.986	93-96%	~90%	NR	NR	>95%
Aggregate Status	Mostly monomer	Mostly monomer	Mostly monomer	Mostly monomer	monomer 79.6%. Dimer 20.4%	93.6% monomer/ 6.4% product-related impurities	NR
Human Serum Recognition	Yes	Yes	Yes	NR	No	No	No
GMP compatibility	Yes	Yes	No	Yes	No	No	No
Animal-Derivative-Free	Yes	Yes	No	No	No	No	No
Variants Tested	Yes	Yes	No	No	Yes	Yes	No
Patent status	Not patented	Not patented	NR	Patented	NR	NR	NR
References	This study	This study	Singh et al 2020	Noe et al 2014	Plassmeyer et al 2009	Plassmeyer et al 2009	Kolodny et al 2001

Note: NR= No reported

Supplementary figures:

A Scale-Up Compatible *Pichia pastoris* Platform for High-Yield Production of Recombinant *Plasmodium falciparum* Circumsporozoite Protein Variants

Neo Judith Lemme^{1,2}, Rampa Victor Mosweu^{1,2}, Ihonga Kuhepa^{1,2}, Amandeep Kaur¹, Shravani Deo¹, Yi-Lin Chen¹, Mei Du¹, Rojelio Mejia³, Zhuyun Liu¹, Wen-Hsiang Chen^{1,3}, Jungsoon Lee¹, Ulrich Strych^{1,3}, Peter Hotez^{1,3,4,5}, Maria Elena Bottazzi^{1,3,4,5}, Nestor L. Uzcategui^{1,3,*}.

¹ Texas Children's Hospital Center for Vaccine Development, Houston, TX, USA

² Ministry of Health Botswana-Government of Botswana, Gaborone Botswana

³ Departments of Pediatrics and Molecular Virology and Microbiology, National School of Tropical Medicine Baylor College of Medicine,
Houston, TX, USA

⁴ Baylor University, Department of Biology, Waco, TX, USA

⁵ Rice University, James A. Baker III Institute for Public Policy, Houston, TX, USA

*Corresponding author: Nestor L. Uzcategui (nestor.uzcateguiaraujo@bcm.edu)

PfaCSP0001

PfaCSP0002

PfaCSP0004

RI cleavage site, N-stabilizing peptide, NANP motif, TSP motif

CIS43: NPDPNANPNVDPNAN **L9:** NANPNVDPNANPNVD

49

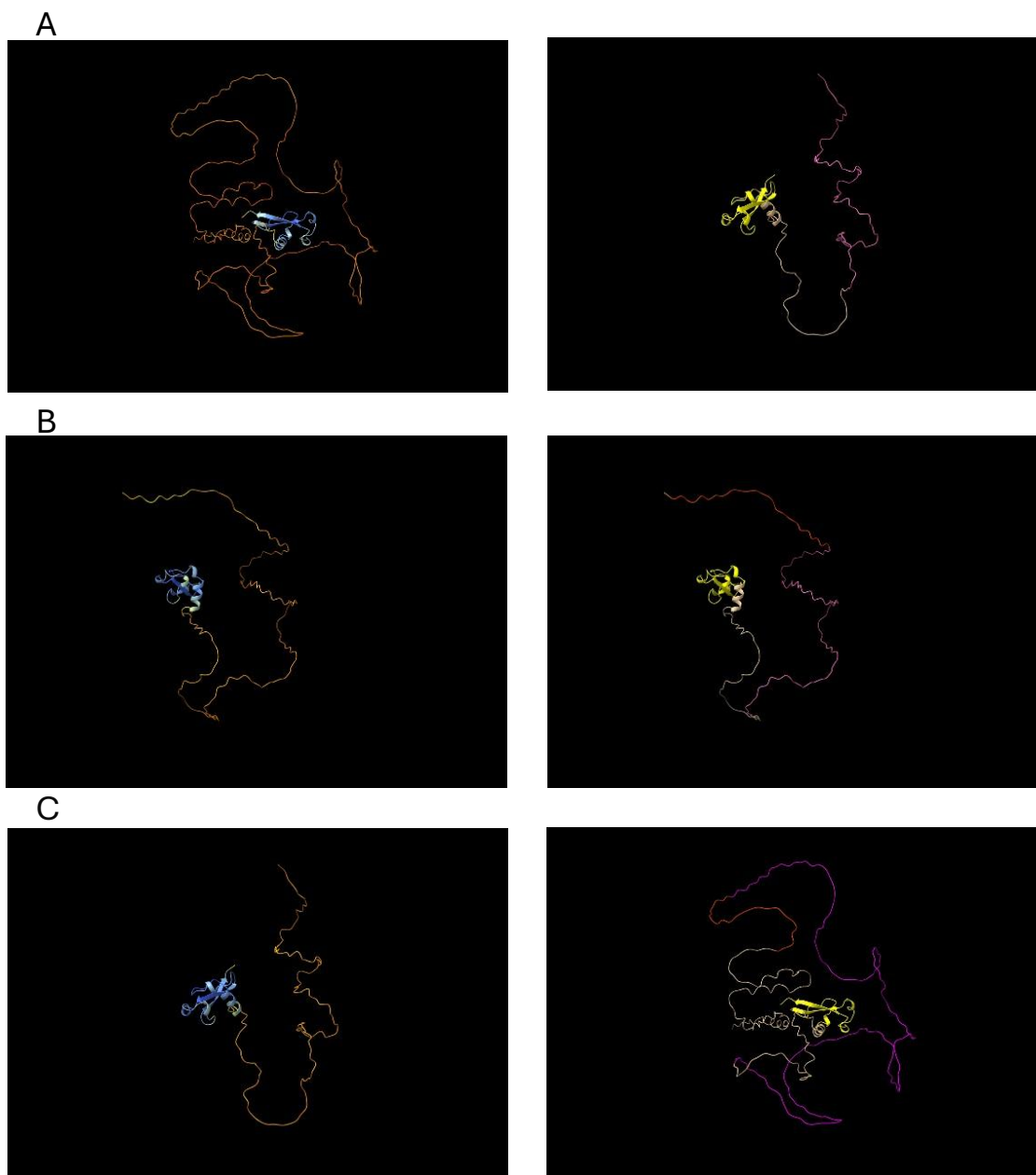


Figure S2. 3D structures generated using AlphaFold 3 for the recombinant proteins **PfaCSP0001 (A)**, **PfaCSP0002 (B)**, and **PfaCSP0004 (C)**. Left panel: predicted structures colored according to pLDDT (predicted Local Distance Difference Test) confidence scores (blue, very high confidence; light blue, confident; yellow, low confidence; orange, very low confidence). Right panel: the same predicted structures colored according to Figure S1 to facilitate identification of the corresponding amino acid regions within the 3D models (orange: N-stabilizing peptide, pink: NANP Repeat motif, yellow: TSP1).

