nature medicine

Article

https://doi.org/10.1038/s41591-023-02659-z

A candidate antibody drug for prevention of malaria

Received: 10 March 2023

Accepted: 20 October 2023

Published online: 2 January 2024

Check for updates

Katherine L. Williams[1] ✉, Steve Guerrero[1], Yevel Flores-Garcia[2], Dongkyoon Kim[1,3], Kevin S. Williamson[1], Christine Siska[4], Pauline Smidt[4], Sofia Z. Jepson[4], Kan Li[5], S. Moses Dennison[5], Shamika Mathis-Torres[2], Xiaomu Chen[1], Ulrike Wille-Reece[6,7], Randall S. MacGill[7], Michael Walker[8], Erik Jongert[9], C. Richter King[7], Christian Ockenhouse[7], Jacob Glanville[10], James E. Moon[11], Jason A. Regules[12], Yann Chong Tan[1,13], Guy Cavet[1,14], Shaun M. Lippow[1], William H. Robinson[15], Sheetij Dutta[12], Georgia D. Tomaras[5,16], Fidel Zavala[2], Randal R. Ketchem[4] & Daniel E. Emerling[1] ✉

Over 75% of malaria-attributable deaths occur in children under the age of 5 years. However, the first malaria vaccine recommended by the World Health Organization (WHO) for pediatric use, RTS,S/AS01 (Mosquirix), has modest efficacy. Complementary strategies, including monoclonal antibodies, will be important in efforts to eradicate malaria. Here we characterize the circulating B cell repertoires of 45 RTS,S/AS01 vaccinees and discover monoclonal antibodies for development as potential therapeutics. We generated >28,000 antibody sequences and tested 481 antibodies for binding activity and 125 antibodies for antimalaria activity in vivo. Through these analyses we identified correlations suggesting that sequences in *Plasmodium falciparum* circumsporozoite protein, the target antigen in RTS,S/AS01, may induce immunodominant antibody responses that limit more protective, but subdominant, responses. Using binding studies, mouse malaria models, biomanufacturing assessments and protein stability assays, we selected AB-000224 and AB-007088 for advancement as a clinical lead and backup. We engineered the variable domains (Fv) of both antibodies to enable low-cost manufacturing at scale for distribution to pediatric populations, in alignment with WHO's preferred product guidelines. The engineered clone with the optimal manufacturing and drug property profile, MAM01, was advanced into clinical development.

Malaria is a mosquito-borne, parasitic disease endemic in regions impacting over 1.5 billion people in Asia, the Americas, the Middle East and Africa. More than 247 million malaria cases and 619,000 malaria-related deaths were reported in 2021 (ref. 1), with 76.8% of these deaths occurring in children under the age of 5 years. WHO has determined that reduction of morbidity and mortality in infants and children due to *Plasmodium falciparum* is among the most urgent priorities in combatting malaria[2]. Although vaccination has been a key tool in control and eradication of other infectious diseases, the development of a vaccine for malaria has been a 50-year challenge[3].

A full list of affiliations appears at the end of the paper. ✉e-mail: kwilliams@atreca.com; danielemerling@biosimplify.com

The first vaccine recommended for use by WHO, RTS,S/AS01 (Mosquirix), targets the circumsporozoite (CSP) protein of *P. falciparum* (PfCSP), the malaria species primarily responsible for mortality in Africa[1,3]. Recently a second vaccine, R21/Matrix M, based on the same CSP-derived antigen as RTS,S[4], was also recommended for use[5]. Immunization with CSP antigen induces anti-CSP antibodies that act by binding to sporozoites, the infective form of the malaria parasite introduced by mosquito bite, and by inhibiting their initial infection of liver cells[6,7]. However, vaccine efficacy against clinical malaria induced by RTS,S/AS01 in children is limited to 45% after the first dose and wanes to 36% over 4 years of follow-up[8]. Chemoprevention-based prophylactics are an alternative to vaccines, but drug resistance and complex drug regimens that lead to poor adherence[1] limit their utility.

Longlasting antimalaria monoclonal antibody (mAb) prophylaxis could complement these existing prevention strategies by providing immediate and more stable serum antibody levels and limiting adherence concerns[2]. However, the potential high cost of antibody drugs can be a barrier to access in low-to-middle-income countries (LMICs) and administration to pediatric populations requires small-volume doses of highly concentrated drugs amenable to low-viscosity formulations for administration via small needles. For these reasons, the WHO guidelines for monoclonal antibody use in malaria prevention recommend early consideration of key factors that contribute to the cost of goods, including manufacturing properties, and to drug viscosity, including biophysical properties such as protein–protein interactions, protein aggregation and protein conformational and colloidal stability[9].

Recent reports show that treatment with mAbs[10] can completely prevent malaria following controlled human infection[11] and provide 88% efficacy (prevention of infection) for 6 months in endemic regions[12], although limited details have been published about the biophysical properties consistent with manufacturing or cost-effective administration of these antibodies to pediatric populations living in LMICs. Two mAbs already tested in clinical trials, L9 (ref. 13) and CIS43 (refs. 11,12), were isolated from B cells of vaccinees immunized with whole sporozoites and can prevent malaria infection by targeting specific epitopes on CSP. CSP comprises three main domains: (1) an N terminus; (2) a central repeat region composed of multiple (25–40) NANP tetrapeptides ('major repeat') interspersed with an NPDP tetrapeptide and two to four NVDP ('minor repeat') tetrapeptides; and (3) a C-terminal domain[14,15]. L9 and CIS43 preferentially bind epitopes containing, respectively, NPNV[16] in the minor-repeat region and DPNA[17] in the 'junctional region' (JR) that links the N-terminal and repeat domains. However, both mAbs can promiscuously bind NPNA epitopes in the central repeat region[16]. These NPNA repeats are conserved across all Pf strains[18,19] and are the only tetrapeptides included in RTS,S. A third mAb, AB-000317 (ref. 20), preferentially binds NPNA epitopes. Despite its potent inhibitory activity in vivo[16,20,21], AB-000317 was not advanced into clinical development due to unpublished evidence of suboptimal cell line expression levels and human tissue cross-reactivity.

Here we report on the discovery of >50 preclinically protective antibody lineages (including the original discovery of AB-000317) by sequencing plasmablast (PB) repertoires from 45 RTS,S/AS01 vaccinees. In conducting these studies we observed that, for many vaccinees, antibodies expressed by these circulating PBs were insufficient to protect against malaria challenge. Subsequent investigation uncovered an inverse association between vaccinee protection status and CSP reactivity of antibodies expressed from their PB repertoires, suggesting that RTS,S/AS01 vaccination may induce immunodominant anti-CSP responses that do not contribute to protective immunity. Given the published support for prophylaxis with mAbs as a strategy against malaria[11–13], and recent advances using mAbs as therapeutics and prophylactics against infectious diseases[22], we aimed to select and engineer (Supplementary Fig. 1) a potent and longlasting antibody drug with biophysical properties amenable for cost-effective manufacturing and dosing in pediatric populations[9,23].

Results

Initial mAb library includes representative clones from dominant PB lineages

Peripheral blood mononuclear cells (PBMCs) were obtained from vaccinees enrolled in a phase 2a clinical trial[24] (NCT01857869) designed to evaluate either three full doses of RTS,S/AS01$_B$ 1 month apart (012M; $n = 15$) or two full doses 1 month apart and a smaller (one-fifth, 'fractional') dose 6 months later (Fx017M, $n = 30$). Vaccinees were challenged with malaria in a controlled human malaria infection (CHMI) model 3 weeks after the third dose. A subset ($n = 19$) received a fourth dose 8 months after the third dose and were subsequently challenged a second time with malaria (Supplementary Fig. 2). PBs were isolated from PBMCs collected 7 days post third (P3D; $n = 22{,}319$ PB) and post fourth doses (P4D; $n = 10{,}629$ PB; Supplementary Table 1) and were used to generate natively paired heavy- and light-chain IgG sequences using our Immune Repertoire Capture sequencing platform. Almost all (99.2%) of the antibody sequences were divergent from inferred germline precursor sequences (Extended Data Fig. 1 and Methods). Consistent with previous malaria studies[17,25–28], specific germline heavy- and light-chain genes and pairings, including IGHV3-30/33, KV1-5, KV3-20 and LV1-40, were observed frequently in the dataset (Supplementary Figs. 3–5). No significant associations were observed between vaccinee protection status or between vaccinee dose group and multiple IgG sequences and repertoire features (Supplementary Figs. 3–7).

P3D and P4D PBs that were probably derived from a common progenitor B cell clone were grouped into Ig lineages ($n = 18{,}980$; Methods). Lineage size ranged from 1 to 84 PBs (P3D) and from 1 to 93 PBs (P4D). Because PBs have a short half-life in blood[29] (reviewed in refs. 30,31) and were isolated from a small volume of blood (~10 ml), detection of lineages with two or more PBs probably indicates recent cellular expansion in lymphoid organs. One-fifth of lineages contained at least two PBs ('expanded lineages', 19.4%, $n = 3{,}684$ lineages; Fig. 1a). Consistent with antigen-driven selection pressure following vaccination, most of these lineages were clonally expanded and included two or more distinct antibody sequences (Fig. 1b). Several lineages had clonal representatives that were observed after both the third and fourth immunization ('recalled lineages'; 4.1–26.6% of vaccinee P3D-expanded lineages). Furthermore, we observed that many of these expanded lineages also showed evidence of sequence convergence between two or more vaccinees (7.3–46.7% of P3D-expanded lineages; Methods). Not surprisingly, lineages with only a single observed PB in P3D repertoires ($n = 10{,}841$) had significantly lower rates of convergence (2.0–13.8%) and recall (1.2–18.6%) than expanded lineages ($P < 0.0001$ and $P < 0.001$, respectively, Wilcoxon matched-pairs, two-tailed test) and had higher levels of somatic hypermutation (SHM; Extended Data Fig. 2a). Thus, to increase the likelihood of identifying antibodies derived against the RTS,S antigen we mainly focused subsequent analyses on expanded lineages (Fig. 1b).

We hypothesized that lineages with the largest number of PBs per vaccinee ('dominant lineages') were more likely to target the vaccine because they had outcompeted other PB lineages for antigen binding and/or T cell help in lymphoid organs. Thus, for each vaccinee, expanded P3D lineages were rank ordered by size ('rank-size'). The sum of PBs in the top four rank-sized lineages constituted 17–100% of each vaccinee's P3D-expanded lineage repertoire and 33% of PBs in all of the vaccinees' P3D-expanded lineages (Fig. 1c). Because this pattern of PB distribution was consistent across vaccinee protection status and dose regimens (Extended Data Fig. 2b,c), we generated a mAb screening library that enriched for the dominant P3D lineages of both protected and unprotected vaccinees.

CSP-reactive mAbs have less SHM and are more predominant in unprotected vaccinees

We selected a library of 369 clones, each representing a unique P3D lineage, for testing (Methods and Supplementary Fig. 8a). This library

https://doi.org/10.1038/s41591-023-02659-z

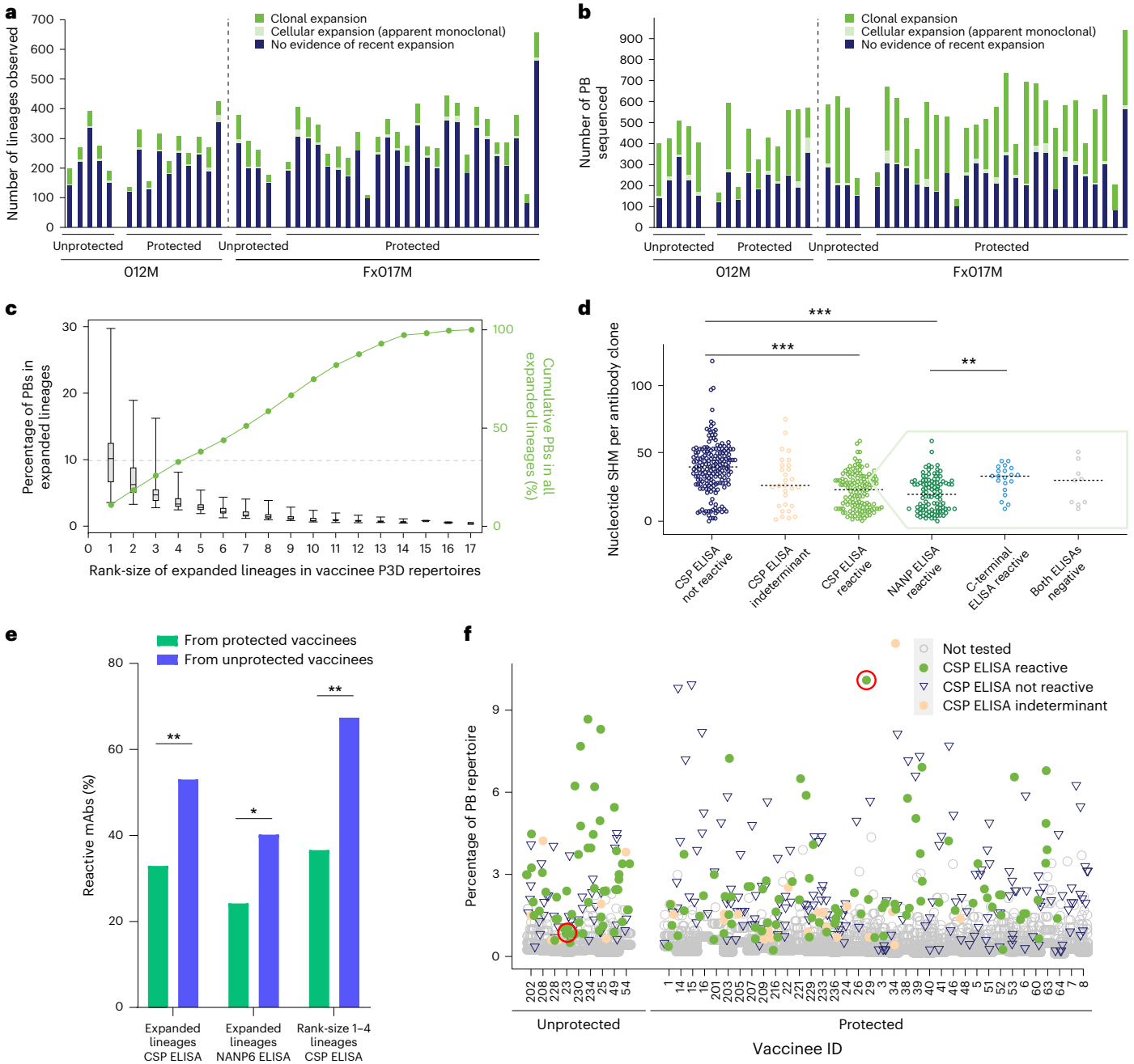


Fig. 1 | CSP-reactive lineages from blood PBs following the third dose of RTS,S. a,**b**, IgG lineages for each vaccinee (bars, $n = 45$) that are clonally expanded (that is, have two or more distinct nucleotide clones; green), that consist of two or more identical nucleotide clones (gray) or that contain only one observed PB (blue) are shown by either number of lineages (**a**) or number of PBs per vaccinee (**b**). **c**, By vaccinee ($n = 45$), the size of each expanded lineage was calculated by dividing the number of PBs in that lineage (2–84 PBs per lineage) by the number in each vaccinee's P3D-expanded lineage repertoire (37–492 PBs per vaccinee) and then assigning a rank-size. Boxes indicate interquartile ranges, lines within boxes are medians and whiskers represent minimum and maximum across vaccinees for each rank-size (44–1,246 PBs per rank, 22–301 lineages per rank). Dotted line indicates that the top four rank-size lineages contain 33% of PBs in all P3D-expanded lineages ($n = 11,478$ PBs, 2,662 lineages). **d**–**e**, ELISA reactivity, SHM levels and vaccinee protection status of mAbs from expanded lineages ($n = 349$). **d**, Number of nucleotide mutations from germline (SHM) for mAbs that are not reactive (dark blue, $n = 185$), show an indeterminant, weak signal (orange, $n = 29$) or are reactive (light green, $n = 135$) in a CSP ELISA. Domain specificity for CSP-reactive mAbs is shown in the box with a dashed green outline. Monoclonal antibodies reactive by ELISA to the NANP6 repeat-region peptide (dark green, $n = 98$), to the C-terminal region peptide (Pfs16, royal blue, $n = 20$) or not reactive in either peptide ELISA (gray, $n = 9$). Lines represent median values, $^{***}P < 0.0001$, $^{**}P < 0.001$, unpaired two-tailed Mann–Whitney test. CSP-reactive mAbs that were not tested in peptide ELISAs ($n = 8$) are not shown. **e**, Percentage of tested antibodies from expanded lineages that originated from protected (green, $n = 36$) and unprotected (blue, $n = 9$) vaccinees that are CSP-reactive (82 out of 249, and 53 out of 100 mAbs, respectively), repeat-region NANP6 peptide-reactive (59 out of 244, and 39 out of 97 mAbs, respectively) and the subset from only dominant rank-size one to four lineages that are CSP-reactive (52 out of 142, and 31 out of 46 mAbs, respectively); $^{**}P < 0.001$, $^{*}P < 0.01$, two-tailed Fisher's exact test. **f**, For each vaccinee (x axis), each symbol indicates a single lineage. The lineages ($n = 369$) from which a clone was selected for testing are indicated by CSP reactivity: CSP-reactive (green dots, $n = 139$), indeterminant (orange dots, $n = 29$) or not reactive (royal blue triangles, $n = 201$). All lineages that were not tested are shown (gray circles, $n = 13,134$; 2,313 expanded and 10,821 single-PB lineages). Protected vaccinees have a lower ratio of CSP-reactive versus nonreactive lineages than unprotected vaccinees (bootstrap analysis, $P = 0.0011$). Red circles indicate the two lineages containing the amino acid sequence of AB-000317.

https://doi.org/10.1038/s41591-023-02659-z

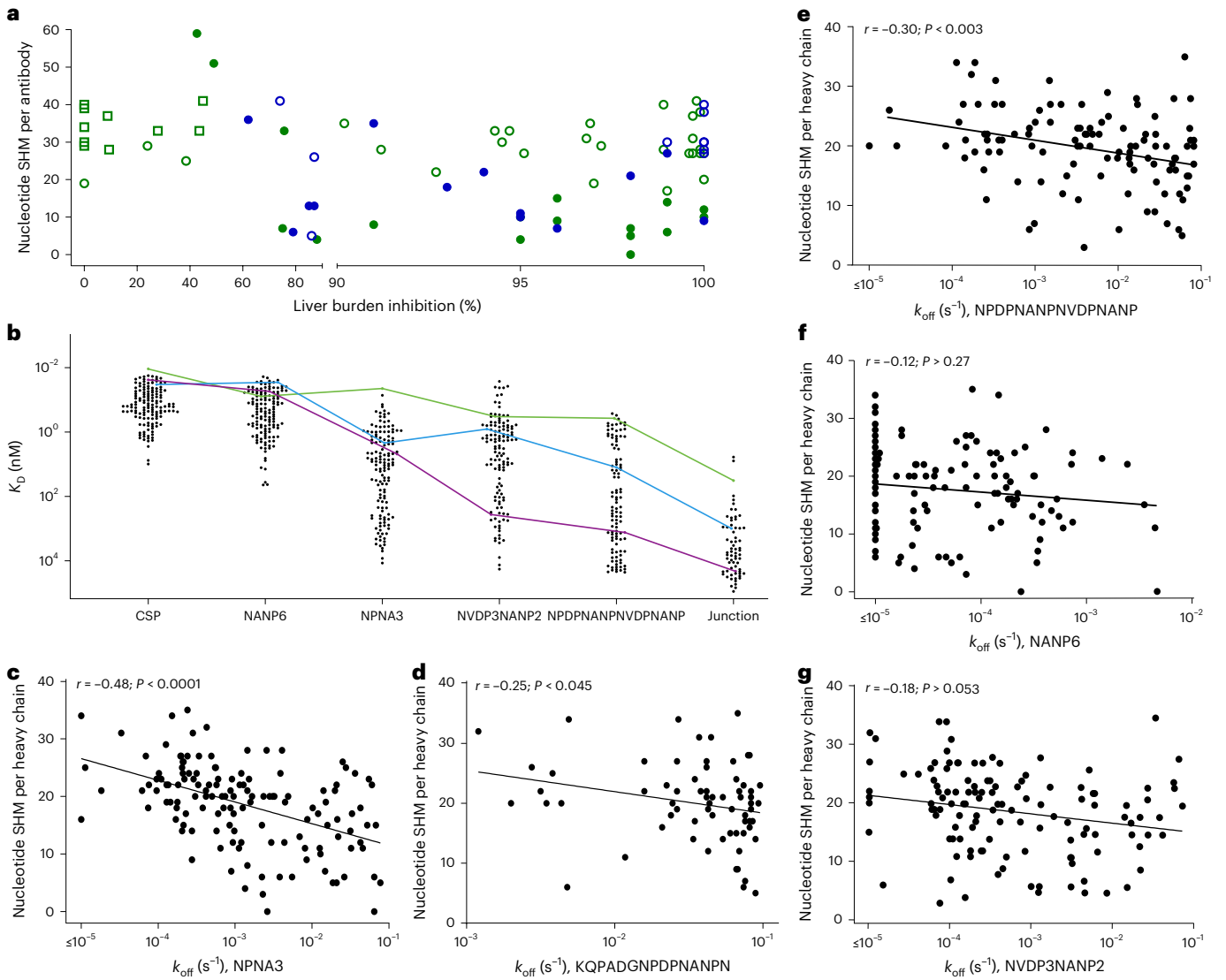


Fig. 2 | Functional mAbs bind CSP-derived peptides not present in RTS,S. **a**, Percentage inhibition in the sporozoite liver burden mouse model and number of nucleotide mutations from germline are shown for mAbs reactive to either NANP6 repeat-region peptide (circles, $n = 67$) and C-terminal-region peptide (squares, $n = 10$), and are indicated as originating from vaccinees who were either protected (green, $n = 54$) or unprotected (blue, $n = 23$) and who received either the standard (012M, closed symbols, $n = 30$) or fractional (Fx017M, open symbols, $n = 47$) dose. **b**, SPR-determined binding potencies (K_d) of mAbs ($n = 141$) selected from 35 of the most efficacious lineages tested against CSP and a panel of CSP-derived peptides that are either homologous (NANP6, NPNA3) or heterologous (NVDP3NANP2, NPDPNANPNVDPNANP, junction [KQPADGNPDPNANPN]) to RTS,S. Examples are shown of a mAb with a broadly promiscuous binding profile (green, AB-007163), another with a profile relatively biased toward homologous peptides (purple, AB-007143) and a third with a profile between these extremes (blue, AB-007175). **c**–**g**, Simple two-tailed linear regression comparing the number of nucleotide mutations from germline (SHM) per heavy chain versus log-transformed SPR binding off-rate (k_{off}) against peptides. **c**, Short, major repeat (NPNA3, $n = 140$). **d**, Junctional (KQPADGNPDPNANPN, $n = 68$). **e**, Short, minor repeat (NPDPNANPNVDPNANP, $n = 109$). **f**, Long, major repeat (NANP6, $n = 141$). **g**, Long, minor repeat (NVDP3NANP2, $n = 129$). See Extended Data Table 1 for correlation analyses with nontransformed data. Dissociation rate (k_{off}) measurements were limited to a minimum of 10^{-5} s^{-1}. MAbs, with rates ≤10^{-5} s^{-1} included in the graphic but excluded from correlation analyses.

included almost all (96%) of the largest lineages (rank-size 1); approximately half (56%) of the second, third and fourth rank-size lineages; a minor subset (6.9%) of smaller rank-sized lineages (rank-size five or more); and a few single-PB lineages (0.18% of 10,841 single-cell lineages). All mAbs ($n = 369$) were screened in a CSP ELISA (Fig. 1d and Supplementary Fig. 8b), and a subset ($n = 130$) were screened against the other RTS,S component, hepatitis B surface antigen (HBsAg; Supplementary Fig. 8c). Of the mAbs screened in both assays, 52% (67 out of 130) were reactive to either CSP ($n = 28$) or HBsAg ($n = 38$) and one additional antibody was reactive in both assays. In total, 38% (139 out of 369) were reactive to CSP and binding for an additional 29 mAbs was indeterminate. Of the CSP-reactive mAbs, 73% (102 out of 139, including AB-000317) bound the NANP6 peptide and 14% (20 out of 139) bound peptides from the C-terminal region (Fig. 1d and Supplementary Table 2).

Given that expanded lineages were more likely to show evidence of convergence and recall compared with single PBs, we tested whether those same features were associated with CSP reactivity. Indeed, mAbs from convergent or recalled lineages were more likely to bind to CSP compared with those from lineages observed in only one vaccinee (54% (55 out of 102) versus 31% (84 out of 267), $P = 0.0001$, two-tailed Fisher's exact test) or with those from lineages that did not show evidence of recall (49% (43 out of 87) versus 19% (16 out of 83), $P < 0.0001$, two-tailed Fisher's exact test). However, recall was significant only after controlling for both factors (adjusted odds

ratio = 3.911, 95% confidence interval (CI): 1.983–8.002; Supplementary Table 3).

Similar to other reports describing immunization with whole sporozoites[27,32], we found that many CSP-reactive mAbs had significantly lower levels of SHM than CSP nonreactive mAbs ($P < 0.0001$; Fig. 1d), even after controlling for potential confounders (adjusted coefficient = −14.22, 95% CI: −17.68 to −10.75; Supplementary Table 3), and SHM levels of NANP6-reactive mAbs were lower than those of C-terminal-binding mAbs ($P < 0.006$; Fig. 1d). We also observed that SHM levels of NANP6-binding mAbs were correlated with dose schedule (Extended Data Fig. 3a) but not with P3D protection status ($P > 0.6$; Extended Data Fig. 3b) as previously reported for all IgG sequences[24].

Surprisingly, the percentages of CSP-reactive ($P < 0.0007$; Fig. 1e) and NANP6-reactive, PB-derived mAbs ($P < 0.006$; Fig. 1e) were lower among P3D-protected vaccinees than P3D-unprotected vaccinees, even after controlling for dose group and convergence (adjusted odds ratio CSP = 0.4109, 95% CI: 0.2512–0.6681; adjusted odds ratio NANP6 = 0.4641, 95% CI 0.2745–0.7845; Supplementary Table 3). This inverse relationship between antigen binding and protection status was also observed when analysis was restricted to only mAbs from the most dominant lineages (rank-size one to four, $P < 0.0004$; Fig. 1e) or when it was extended to include the 20 mAbs from lineages with only one PB ($P < 0.0005$, Fisher's exact test and $P = 0.001$ by bootstrap analysis; Fig. 1f). These data suggest that simply having higher quantities of circulating PBs expressing CSP- or NANP-reactive antibodies may be insufficient for protection from CHMI.

Sporozoite-inhibitory antibodies in P3D PBs are not sufficient for P3D protection

Given the well-reported protective activity of CSP-binding mAbs in both humans[11,13] and mice[17,20,25,26,33,34] and the surprising inverse association between CSP reactivity and vaccinee protection status, we selected 77 mAbs (Methods) from both protected and unprotected vaccinees for advancement in vivo[21,35]. Over half of these mAbs (44 out of 77) provided ≥95% inhibition of sporozoite liver burden, with some demonstrating near-complete protection (≥99.9% inhibition) in C57BL/6 mice ($n = 10$) infected intravenously with chimeric *Plasmodium berghei* (Pb) sporozoites expressing PfCSP rather than PbCSP. All 44 mAbs bound the NANP-repeat region of CSP, with most being derived from the IGHV3-33 germline. Thirteen other NANP-binding, IGHV3-30/33 mAbs demonstrated limited inhibition of parasite liver burden (80–95%) and 12 mAbs, including three C-terminal peptide binders, showed minimal inhibition (20–80%; Fig. 2a and Supplementary Table 2).

Roughly one-third of mAbs tested in vivo (30%, 23 out of 77) were from unprotected vaccinees, including half (7 out of 14) that showed near-complete protection in mice (≥99.9% inhibition; Fig. 2a and Supplementary Table 2), but neither P3D protection status nor vaccinee dose regimen remained significantly correlated with mouse liver burden inhibition in a multivariable linear regression model (Supplementary Table 3). These data suggest that PB expression of these inhibitory antibodies alone is insufficient to drive protection against CHMI. For example, the highly effective antibody AB-000317 (ref. 20) was observed in both a protected and an unprotected vaccinee (Fig. 1f, red circles). However, the AB-000317 lineage was the largest PB lineage in the protected vaccinee but the seventh rank-size lineage in the unprotected vaccinee (respectively, 15.8 versus 1.7% of PB in expanded P3D lineages,). While anecdotal, these data are consistent with the hypothesis that antibody-mediated protection induced following RTS,S is driven by a confluence of immunological factors including, but not limited to, B cell receptor recombination and affinity maturation, as well as other factors not measured here including anti-CSP titer in blood, the contribution of other B cell and T cell subsets and/or other immune factors[36].

Inhibitory antibodies from vaccinees bind CSP peptides not included in RTS,S

To explore the developability of these inhibitory mAbs as potential drugs, 35 NANP-repeat-binding lineages originating from CHMI-protected vaccinees were selected for further pharmacology studies from the 52 that demonstrated ≥90% inhibition in the sporozoite-challenge screen (Methods). To avoid sequence features that can potentially limit drug developability, and to survey lineages with extensive clonal diversity, multiple clones were chosen from 23 of the 35 lineages. Overall, 141 mAbs from 21 protected vaccinees representing both RTS,S dose regimens and a range of SHM levels were selected for testing.

These downselected antibodies displayed a broad range of affinities against CSP as measured by surface plasmon resonance (SPR) (CSP-binding affinity, K_d, of 11 pM–9.8 nM; Fig. 2b and Supplementary Table 4) and demonstrated a significant association between K_d and nucleotide SHM levels ($P < 0.004$, $r = -0.27$ and $P < 0.0003$, $r = -0.34$ for heavy and light chain, respectively, Spearman test), consistent with affinity maturation to CSP that occurred following RTS,S. Correlations between K_d and heavy- and light-chain SHM levels are probably driven by the binding association rate (k_{on}) to CSP ($P < 0.0005$, $r = -0.29$, Spearman) because the correlations between dissociation rates (k_{off}) and SHM levels were not significant (Extended Data Fig. 4a,b and Extended Data Table 1).

Antibodies were also evaluated by SPR for binding to short (12–15 residues) and long (20–24 residues) peptides derived from the varied tetrapeptide-based homologous (NPNA3 and NANP6 peptides) and heterologous (NPDPNANPNVDPNANP, NVDP3NANP2) and junctional [KQPADGNPDPNANPN]) peptides of CSP[16,17,20,26,28,34,37,38] (Fig. 2b). Among the strongest correlations we observed were inverse relationships between SHM and k_{off} to the short major-repeat peptide, the short minor-repeat peptide and the JR peptide (Fig. 2c–e and Extended Data Table 1). k_{off} rates calculated against the long homologous and heterologous peptides and against CSP either show weaker, but still statistically significant, correlations with SHM levels or altogether insignificant correlations (Fig. 2f–g and Extended Data Table 1). This was unexpected given that the long version of the homologous peptide and CSP contain more repeats of the same epitopes present in the short peptide. However, it could be that the multiple tetramers in the NPNA6 peptide may allow for certain antibodies to bind via multivalent avidity[20,39–42], resulting in slow dissociation with k_{off} estimates that were limited to 10^{-5} s^{-1}, the lowest k_{off} to be accurately determined during data analysis. Because these antibodies were excluded from correlation analyses (Fig. 2, Extended Data Table 1 and Methods), we may have underestimated the relationship between SHM levels and avidity to the long NANP peptide. Taken together, these binding data are consistent with reports that some anti-CSP protective mAbs display promiscuous binding across distinct CSP epitopes[16,20,26,28,34,37,43,44]. Thus, we hypothesized that maturation of these inhibitory mAbs to short homologous peptide sequences may also have benefited binding to heterologous peptides.

Anti-sporozoite activity correlates with CSP-peptide binding and SHM levels

Seventy mAbs, representing 33 of the 35 protective lineages evaluated in binding studies, were directly compared in a sporozoite-challenge mouse model with the highly efficacious mAb AB-000317 (refs. 20,21,28,34,38,40,41,45) to prioritize mAbs for development. Antibodies inhibited 44.1–97.5% of sporozoite liver burden (47.4–103.8% of AB-000317 inhibition; Supplementary Table 4). Overall, about one-half of the mAbs demonstrated inhibition comparable to AB-000317 ($n = 32$) while the other half demonstrated significantly weaker inhibition ($n = 36$). One antibody, AB-000224, demonstrated superior activity to AB-000317 (Fig. 3a,b and Supplementary Table 4). Serum concentrations in mice for most mAbs were at least 1,000-fold

https://doi.org/10.1038/s41591-023-02659-z

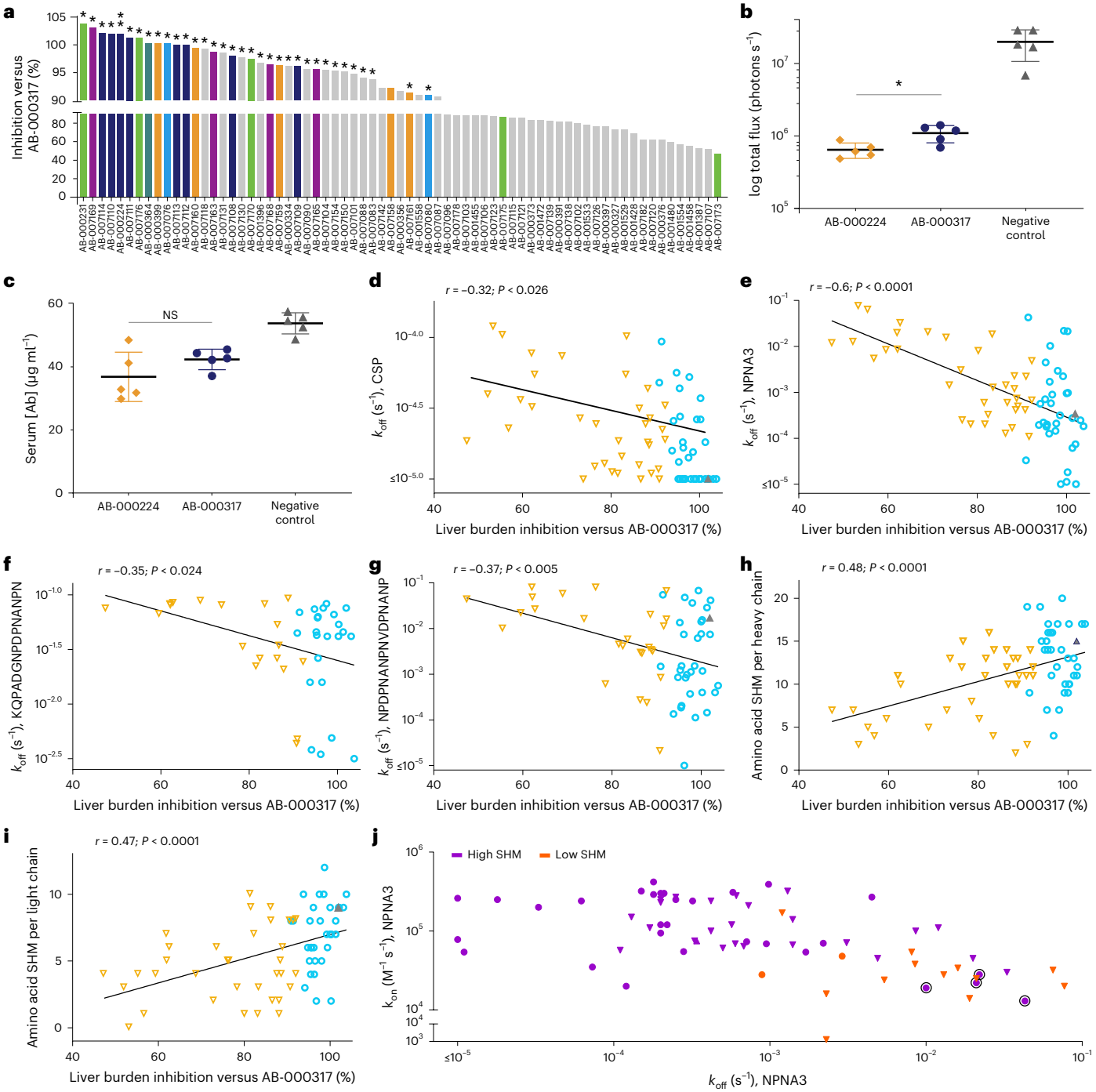


Fig. 3 | In vitro binding, in vivo activity and SHM associated with developable mAbs. a–c, Liver burden model. **a**, Percentage inhibition compared with untreated, infected mice (geometric mean, $n = 5$, 100 μg per mouse) normalized to the activity of AB-000317 of $n = 69$ antibodies (32 lineages, $^{*}P > 0.05$, $^{**}P < 0.05$ activity greater than AB-000317, no icon $P < 0.05$ activity less than AB-000317, two-tailed, nonparametric log-rank), with colors other than gray indicating the six lineages containing the most efficacious mAbs. **b**,**c**, Example data from AB-000224 and AB-000317 of parasite bioluminescence in the liver (total flux, photons s^{-1}), $^{*}P = 0.03$ (**b**) and serum concentrations (serum(Ab), μg ml^{-1}) of mAb at the time of sporozoite challenge (**c**), $P > 0.2$ (NS), mean ± s.d. ($n = 5$ mice), two-tailed Mann–Whitney test. **d–i**, Simple two-tailed linear regression of percentage liver burden inhibitory activity compared with untreated, infected mice (geometric mean, $n = 5$) and normalized to the activity of AB-000317, with each mAb indicated as having activity either significantly better (gray upward-pointing triangle), not different to (cyan circles) or weaker than (orange downward-pointing triangles) AB-000317 (two-tailed, nonparametric log-rank) versus **d–g**, log-transformed binding off-rate (k_{off}) against CSP ($n = 70$) (**d**), major repeat (NPNA3, $n = 70$) (**e**), junctional (KQPADGNPDPNANPN, $n = 42$) (**f**) and short minor-repeat (NPDPNANPNVDPNANP, $n = 60$) peptides (**g**) and versus **h**,**i**, the number of amino acid changes from germline (SHM) for each mAb ($n = 69$)—heavy (**h**) and light chain (**i**) (see Extended Data Table 1 for correlation analyses of non-log-transformed data). Dissociation rate (k_{off}) measurements were limited to a minimum of 10^{-5} s^{-1} and mAbs were excluded from correlation analyses if assigned this value. **j**, log-transformed SPR binding off- (k_{off}) versus on-rates (k_{on}) against NPNA3 peptide of mAbs with either high SHM (purple, ≥20 nucleotide mutations per clone, $n = 55$) or low SHM (orange, <20 nucleotide mutations per clone, $n = 14$) and with activity weaker (downward-pointing triangles), not different to (circles) or better than (upward-pointing triangle) AB-000317 (two-tailed, nonparametric log-rank). mAbs from a lineage reported to bind CSP with Fab–Fab homotypic interactions are indicated (black circles, AB-000399 (refs. 41,42), AB-007159, AB-007160, AB-007161).

https://doi.org/10.1038/s41591-023-02659-z

higher than the CSP K_d of the respective mAbs (Supplementary Table 4 and Fig. 3c), indicating that antibodies demonstrating weak inhibition were not likely to be due to low levels of circulating antibody. Lineages with at least one mAb that demonstrated activity consistent with AB-000317 were considered for further advancement.

To determine whether RTS,S-driven affinity maturation may have contributed to in vivo activity in this second set of downselected mAbs, we assessed whether peptide-binding kinetics or SHM levels correlated with percentage inhibition in the liver burden model. Inhibitory activity was associated with slower k_{off} rates from CSP (Fig. 3d), from the short homologous peptide, (Fig. 3e), from the heterologous JR (Fig. 3f) and from the short (Fig. 3g) and long minor-repeat ($P < 0.002$, Spearman; $P > 0.3$, Pearson) peptides (Extended Data Table 1). Consistent with the hypothesis that maturation to short NANP sequences may induce more inhibitory activity than multiple repeats of NPNA-based epitopes[37,46], we did not observe significant correlations between inhibitory activity and binding kinetics with the long homologous peptide NANP6 ($P > 0.3$ (k_{off}); $P > 0.7$ (k_{on}), Spearman and Pearson, respectively; Extended Data Table 1). Further, the total number of nucleotide and amino acid changes from germline are significantly correlated to percentage inhibition in the liver burden model, where mutational burden was treated as either a continuous (Fig. 3h,i and Extended Data Table 1) or a categorical variable. In the latter model, antibodies with lower mutational burden (86% (12 out of 14) mAbs with <20 nucleotide mutations) demonstrated significantly weaker inhibition compared with AB-000317 than more mutated mAbs (44% (24 out of 55) mAbs with ≥20 mutations, $P = 0.007$; Fisher's exact test, two-tailed). Taken together, these correlations suggest that maturation to NANP-based epitopes in RTS,S may have a bystander affect that enhances binding to structurally similar but heterologous sequences and may be functionally important[16,20,26].

Despite the correlations among SHM levels, binding kinetics and inhibitory activity, some mAbs with high SHM levels are exceptions. In some cases, high-SHM mAbs have poor binding kinetics (slower than median k_{on} values and faster than median k_{off} values in comparison with the short homologous peptide) and are poor inhibitors compared with AB-000317 (Fig. 3j; 23% (16 out of 69)). These antibodies may have resulted from inefficient affinity maturation and/or aberrant selection mechanisms related to survival in and/or recall from memory[32,37,47–49] (Fig. 1e–f). In other cases, some high-SHM mAbs have similarly unfavorable binding kinetics but still demonstrate strong inhibitory activity (Fig. 3j; 12% (8 out of 69)). In these latter cases, affinity maturation toward antibody homotypic Fab–Fab interactions may contribute to anti-CSP-binding potency and increased functional activity[27,39,50], and have been reported for some of the mAbs described here[39,41,42]. Such homotypic interactions may not be reflected in binding kinetics to the short NPNA3 peptide which, due to its short length, cannot sterically accommodate multiple simultaneous binding events[27,39,42].

Indeed, four of the eight mAbs that have unfavorable binding kinetics but are comparable to AB-000317 in activity originate from a lineage containing a mAb that binds via Fab–Fab homotypic interactions (AB-000399 (refs. 41,42); Fig. 3j, red circles). Overall, these data are consistent with mAb affinity maturation via multiple different modes of binding[20,39–42] and reveal >30 mAbs with activity comparable to that of AB-000317 and the potential for development to clinical leads.

Lead antibodies prioritized for development

To identify the most optimal lineage(s) for clinical development, we compared mAbs in regard to the pharmacological and biophysical characteristics necessary for successful manufacture and formulation[9]. Using the sporozoite liver burden data, we further downselected to 26 mAbs representing 15 lineages for evaluation in the parasitemia challenge model, in which mosquitoes infected with chimeric Pb parasites encoding full-length PfCSP fed on C57Bl/6 mice ($n = 10$). Blood-stage

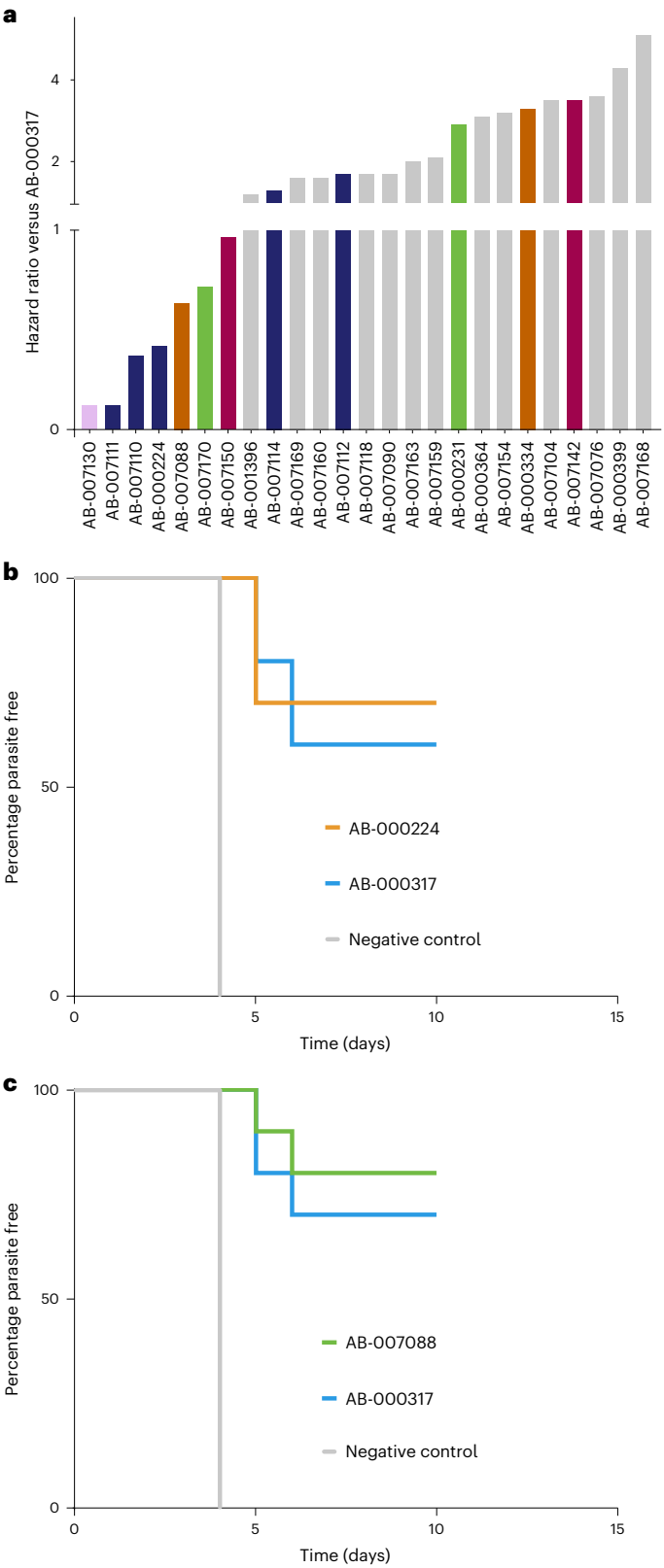


Fig. 4 | Inhibition of mosquito-bite parasitemia from mAbs prioritized for development. a, Hazard ratios of $n = 25$ antibodies (14 lineages) compared with AB-000317 in the mosquito-bite parasitemia model ($n = 10$ mice, 150 µg per mouse), with colors other than gray indicating the five lineages containing the most efficacious antibodies. **b**,**c**, Survival curves from repeat experiments in comparison with AB-000317 and AB-000224 (0.74 (0.15, 3.8)) (**b**) and AB-007088 (0.61 (0.097, 3.8)) (**c**); $n = 10$ mice, two-tailed, nonparametric log-rank (Mantel–Haenszel hazard ratio (95% CIs)).

https://doi.org/10.1038/s41591-023-02659-z

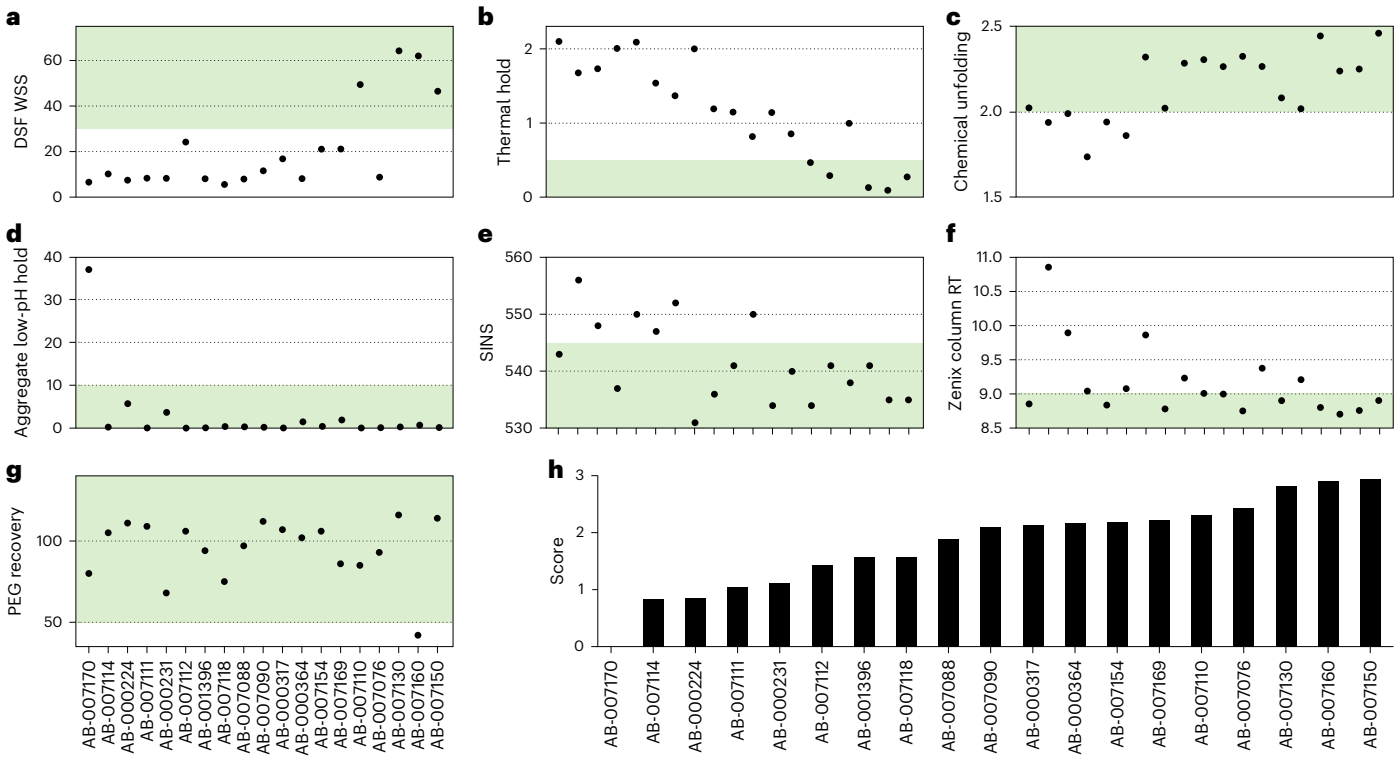


Fig. 5 | Developability properties of lead antibodies. a–g, Developability properties of prioritized lead antibodies as characterized by scores for DSF WSS (**a**), thermal hold (**b**), chemically induced unfolding (**c**), low-pH stability (aggregate low-pH hold) (**d**), SINS (**e**), stand-up monolayer affinity chromatography (Zenix column RT) (**f**) and relative solubility analysis (PEG recovery) (**g**); green shading indicates preferred ranges. **h**, Relative aggregate scores generated from assay panel results.

malaria infection was determined by blood smear on days 4–10 (refs. 21,35). This set of mAbs included AB-000317, AB-000224, 23 other mAbs with liver burden inhibitory activity similar to AB-000317 and one mAb with activity weaker than AB-000317. All except two mAbs were significantly more likely to prevent parasitemia than the negative control. Seven mAbs, including AB-000224 and two other mAbs from the same lineage, displayed a trend towards superior protection versus AB-000317 (nonparametric log-rank hazard ratios <1 versus AB-000317; Fig. 4a and Supplementary Table 4). Serum concentrations in mice for almost all mAbs (25 out of 26) at the time of infection were at least 1,000-fold higher than the respective K_d CSP–SPR of the mAbs (Supplementary Table 4), indicating that mAbs more efficacious than AB-000317 were probably not missed due to low levels of circulating antibody. Overall, the data suggest that AB-000317 and AB-000224 have in vivo activity at or near maximal efficacy among this set of lead antibodies. Given the previous developability concerns with AB-000317, we selected AB-000224 as the lead molecule.

While functional potency is essential for any effective drug, biophysical properties including manufacturability, stability and formulation are equally critical for successful drug development[51,52]. Thus we assessed selected drug properties (for example, protein conformational and solution colloidal stability; Methods) for AB-000224 and 18 other mAbs representing 14 lineages with in vivo activity comparable to that of AB-000317 (Fig. 5). Although none of the data indicate that any of the leads should be eliminated due to development risks, the prime lead (AB-000224) and its siblings demonstrated less favorable characteristics than other mAbs, including relatively poor biophysical stability and aggregation properties (Fig. 5). Thus we selected AB-007088, which demonstrated a similar trend in superiority over AB-000317 in a repeated parasitemia challenge experiment (Fig. 4c and Supplementary Table 4), as the backup molecule given its more favorable developability characteristics (Fig. 5).

Clinical candidate engineered for optimized developability

To address developability risks, we engineered AB-000224 and AB-007088 by mutation of specific residues, primarily in the antibody framework regions, according to the Just—Evotec Biologics' Abacus design platform to avoid impacting complementarity-determining region (CDR)-mediated binding activity (Fig. 6a,i, Extended Data Fig. 5a,i and Methods). Specifically, outlier amino acid resides in framework regions across germline genes were identified via computational covariance evaluation[53] of structurally aligned residue positions. These amino acids were replaced with mutations known to impact molecular stability[54], which often translates to improved developability. For AB-000224 and AB-007088, a total of 17 and five clonal variants, respectively, were designed and tested in the same biophysical and pharmacologic assays used previously. Engineered mutations improved both conformational and colloidal stability of many variants, including enhanced thermal stability (assessed by differential scanning fluorimetry (DSF) and thermal fold assays), solubility and aggregation profiles during storage (assessed in the stand-up monolayer affinity chromatography assay; Fig. 6b–i and Extended Data Fig. 5b–i). Most of the variants retained parental mAb binding profiles against a subset of tested peptides (NPNA3 and NVDP3NANP2; Extended Data Table 2), and activity was not significantly different from AB-000317 for the subset tested in vivo. Unlike the previous screening result that showed AB-000224 to be more efficacious than AB-000317 in the liver burden model (Fig. 3b), average percentage inhibition for AB-000224 and most AB-000224 variants was comparable to AB-000317 (Extended Data Fig. 6 and Extended Data Table 3) despite lower sera antibody levels in some cases (Extended Data Table 3).

Variants were generated as human IgG1 with an Fc mutation that extends antibody half-life (Xtend[55]) and which should compensate for the lower serum antibody levels observed in the mouse studies. These variants were used to make a panel of stable transfectant cells,

https://doi.org/10.1038/s41591-023-02659-z

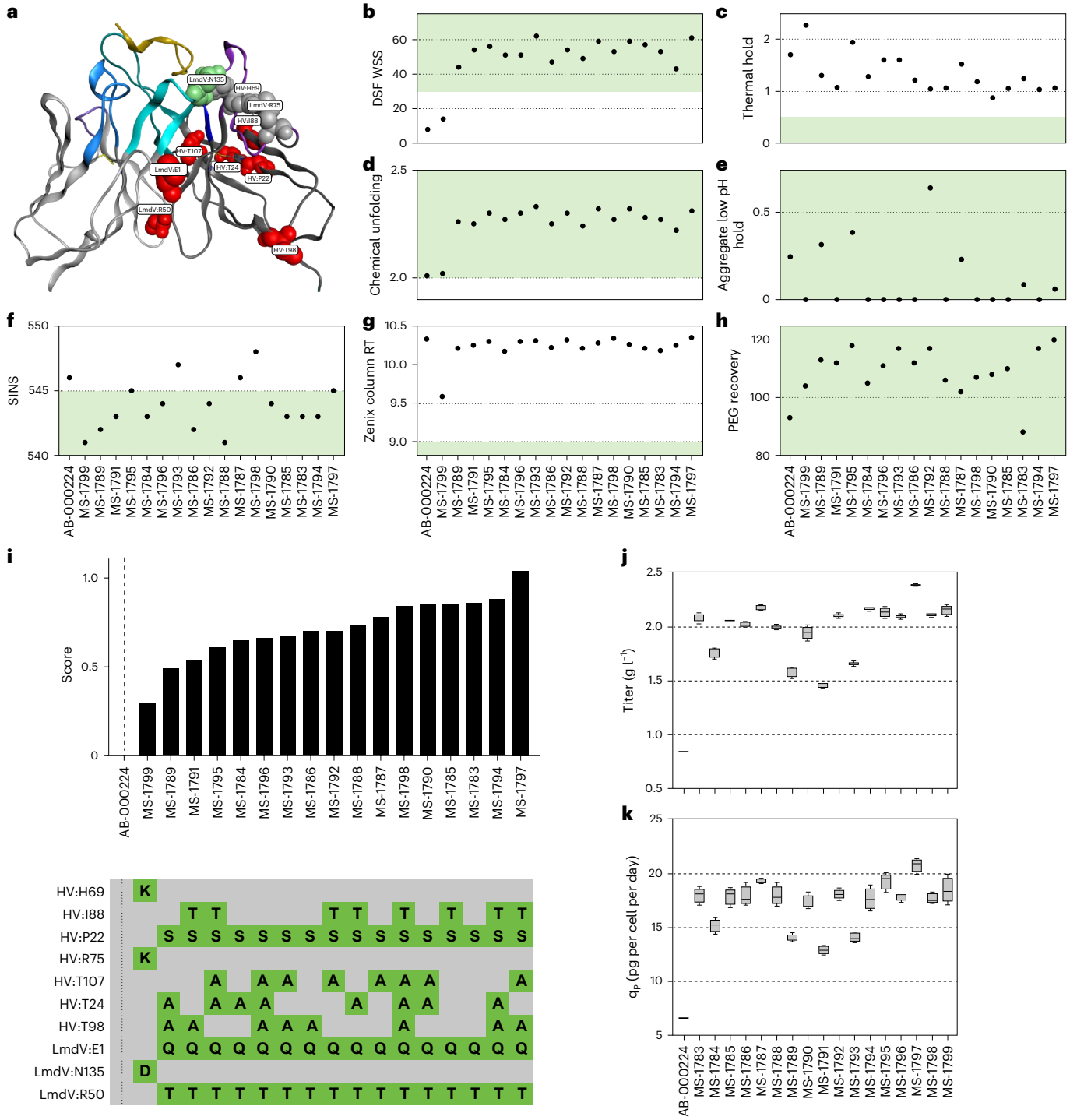


Fig. 6 | Developability properties of engineered variants from selected lead AB-000224. a, Mutations made to generate engineered variants depicted in the Fv region of the Fab of AB-000224 in complex with NPNA4 (gold ribbon, PDB ID 6WFY; Methods). The light-chain framework is shown in silver and the heavy-chain in pale gray. The CDRs for light and heavy chains are indicated, respectively, as CDR1 (light and dark blue), CDR2 (light and dark purple) and CDR3 (light and dark cyan). Mutation sites of engineered variants are labeled as stability violations (red), deamidation sites (green) and paired sibling sites (gray; Methods). **b**–**h**, Developability properties of engineered variants of AB-000224 as characterized by scores from DSF WSS (**b**), thermal hold (**c**), chemically induced unfolding (**d**), low-pH stability (aggregate low-pH hold) (**e**), SINS (**f**), stand-up monolayer affinity chromatography (Zenix column RT) (**g**) and relative solubility analysis (PEG recovery) (**h**); green shading tindicates preferred ranges. **i**, Aggregate score of assay panel results from **b**–**h** for each engineered variant, with mutations shown in comparison with parental AB-000224. **j**,**k**, Characterization of stable cell production pools generated from engineered variants compared with AB-000224 for production titers (g l^{-1}) (**j**) and cell-specific productivity (q_P, pg per cell per day) (**k**). Boxes indicate interquartile ranges, lines within boxes are medians and whiskers represent minimum and maximum (four independent replicates of each antibody).

and antibody production data were collected (Fig. 6j,k and Extended Data Fig. 5j,k) to identify optimal pools for the generation of a production cell line (CHO-K1 derived; Methods). Optimization of stability violations in AB-000224 greatly improved production titers (Fig. 6j,k) which, importantly, can reduce cost per dose in alignment with WHO guidelines[9]. Three cell lines from the panel of 22 variants were selected

for clone generation. The best-producing clonal cell line identified in continuous-perfusion bioreactors was from engineered mAb MS-1797 (Fig. 6j,k). Based on this outcome and its superior scores in developability assays (Fig. 6b–i), MS-1797, renamed MAM01, advanced into production following good manufacturing practices. Material generated from this line is being used to support studies for clinical development (NCT05891236) of an antimalaria drug suitable for use in pediatric populations living in LMICs.

Discussion

Using single-cell sequencing of B cells from RTS,S vaccinees, we generated a library of CSP-specific mAbs that could be assessed and downselected to those most amenable for engineering and development as antimalaria medicines. In doing so we also uncovered important characteristics of the humoral response to RTS,S that may underlie the efficacy and durability of this vaccine.

First, we unexpectedly discovered an inverse relationship between the percentage of CSP-specific, IgG-expressing PBs and vaccinee protection status P3D (Fig. 1e–f). These data suggest that, despite the well-reported association between anti-CSP antibodies and protection following RTS,S[56–60], more cells expressing anti-CSP or anti-NANP antibodies may not drive stronger protection. Because antibodies targeting repeat regions demonstrate a range of sporozoite-inhibitory activity in vivo, the difference between protective and unprotective antibody repertoires in humans could be driven by the relative proportion of highly versus weakly effective repeat-binding antibodies. This point is exemplified by the highly efficacious mAb, AB-000317, which was expressed in the most dominant P3D lineage of a protected vaccinee and in a much less frequent lineage of an unprotected vaccinee (Fig. 1f). Competition between antibodies at the sporozoite surface ('epitope masking')[45,47,48] and/or within lymphoid organs[32], where dominant but weakly functional antibodies outcompete subdominant, highly effective antibodies for binding to the repeat regions, could be one explanation for this inverse relationship between protection and prevalence of repeat-binding lineages (Fig. 1e). Indeed, our data indicating that P3D PB repeat-binding mAbs have lower levels of SHM than other mAbs (Fig. 1d) support the hypothesis that immature clones are preferentially activated and expanded over more protective memory clones[32,48,49]. Furthermore, this hypothesis could underlie other observations about RTS,S responses: specifically, functional antibodies in sera were higher post second dose (P2D) versus P3D[56] and anti-CSP P2D, but not P3D, PB and memory B cells associate with P3D protection status[61] and some vaccinees lost previous protective immune signatures P3D[61]. Overall, we propose that expression of potent, inhibitory antibodies by P3D PBs alone is insufficient for protection. Rather, relative levels of effective versus ineffective repeat-binding antibodies may be important in provision of consistent protection. Together, these data highlight the need for studies to correlate protection with the ratio of effective to ineffective repeat-binding antibodies circulating in sera at the time of infection.

Second, we found that sporozoite-inhibitory activity in mice does not correlate with binding kinetics to the long NANP6 peptide (Extended Data Table 1) but does significantly correlate with k_{off} to CSP and with binding kinetics to both the short NANP-containing peptide (NPNA3) and peptides from the minor-repeat region and JR (Fig. 3d–g and Extended Data Table 1). These data suggest that protective antibodies induced following RTS,S vaccination probably affinity mature to short NANP repeats. Because the short PNANPN sequence is contained within both major repeats (included in RTS,S) and the minor repeat and JRs (not included in RTS,S), affinity maturation against this sequence may allow these antibodies to gain and/or improve promiscuous binding activity to epitopes containing PNANPN or a portion thereof, that may be important for potency[28,39,44] but which are not fully represented in RTS,S. Consistent with this interpretation, aggregate levels of SHM in both heavy and light chains of inhibitory mAbs are correlated with binding kinetics to NANP-, NVDP- and NPDP-containing short peptides (Fig. 2c–e and Extended Data Table 1), as well as with inhibitory activity in the sporozoite-challenge model (Fig. 3h–i and Extended Data Table 1). Overall, the data are consistent with suggestions that next-generation anti-CSP vaccines contain fewer NANP repeats[37,62] and/or should include sequences from minor repeats and JRs[16,28,34,46,63–66].

Last, multiple observations including (1) the inverse relationship between the percentage of CSP- and NANP6-binding antibodies from expanded linages and protection against CHMI (Fig. 1e); (2) the correlation between binding kinetics with NVDP- and NPDP-containing peptides absent in RTS,S (Fig. 3f–g and Extended Data Table 1) and sporozoite inhibition in vivo; and (3) the correlation between binding kinetics to the short NPNA3, but not the longer NANP6 peptide, and sporozoite inhibition in vivo (Fig. 3e and Extended Data Table 1), are consistent with a hypothesis where multiple NANP repeats act as an immune 'decoy'[37,67] or 'smokescreen'[66,68] that dilutes protective immunity[69–73]. Under this hypothesis, antibody lineages that bind only to homologous epitopes are preferentially expanded over promiscuous mAbs that bind well to both homologous and heterologous epitopes. Because many of the antibodies that bind only to NANP repeats offer limited protection in vivo (Extended Data Table 1), enrichment for these antibodies dilutes the protective capacity of the broader anti-CSP repertoire[32,37,48]. In contrast, promiscuously binding antibodies that can simultaneously bind to multiple[39,42,50] NANP repeats and heterologous epitopes may drive superior protection in vivo because they can saturate binding sites and further stabilize[39,42,50,74,75] mAb interactions with both homologous and heterologous epitopes. Consistent with this idea, three of the most active in vivo mAbs, CIS43, L9 and AB-000317, can bind CSP with high stoichimetries[16,20,39] via two binding events[16]. Further studies that examine the mechanisms of mAb binding to junctional, minor- and major-repeat regions and the interaction with sporozoite inhibition are needed to test this hypothesis.

Our study has a number of limitations. First, we did not assess whether the antibodies we characterized from expanded PB lineages collected 1 week before infection accurately reflected the composition of sera antibodies at the time of infection. Second, the current, WHO-recommended RTS,S vaccine includes the adjuvant AS01E[76]. Our study, however, examined antibodies derived from vaccinees who received RTS,S AS01B[24]. While the two adjuvants contain the same components, they are included at different levels[77]. More work will be needed to assess whether RTS,S/AS01E as used in the field produces antibody repertoires like those characterized here. Third, the correlations we observed between binding off-rates and function (Fig. 3d–i and Extended Data Table 1) were limited to mAbs from protected vaccinees. Further studies will be needed to assess whether similar correlations exist for inhibitory, repeat-binding antibodies derived from unprotected vaccinees. Finally, we tested only a small fraction of the sequences we generated. While we focused our discovery campaign on the larger PB lineages from each vaccinee, 3,334 smaller but expanded lineages remain uncharacterized.

By sequencing PBs, which represent the breadth of Ig sequence diversity that originates from lymphoid reactions following RTS,S vaccination, we deconstructed a part of the humoral response from protected and unprotected vaccinees. We identified lineages with highly protective antibodies in mouse models, screened sequence-diverse clones within those lineages for development-related properties and further engineered a clone to optimize its developability characteristics. These properties will increase the likelihood that regimens can be successfully developed for pediatric populations, which require small-volume, concentrated doses at low viscosity[9]. Given that the in vivo efficacy displayed by MAM01 is comparable to AB-000317, which in turn has activity comparable to or better than that of CIS43 (refs. 16,34), we believe that MAM01 will be useful for individuals living in malaria-naïve and -endemic regions and may also meet the WHO's preferred product profile[9], including cost-effective dosing for delivery

in LMICs. By focusing on properties critical for manufacture and distribution to global pediatric populations[9,23], in addition to the requirement for functional potency, the work reported here may contribute to prophylactic strategies that aid efforts in the eradication of malaria.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-023-02659-z.

References

1. World Health Organization. World malaria report 2022. https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2022 (2022).
2. World Health Organization. Global technical strategy for malaria 2016-2030, 2021 update. www.who.int/publications/i/item/9789240031357 (2021).
3. Sinnis, P. & Fidock, D. A. The RTS,S vaccine—a chance to regain the upper hand against malaria? *Cell* **185**, 750–754 (2022).
4. Collins, K. A., Snaith, R., Cottingham, M. G., Gilbert, S. C. & Hill, A. V. S. Enhancing protective immunity to malaria with a highly immunogenic virus-like particle vaccine. *Sci. Rep.* **7**, 46621 (2017).
5. World Health Organization. WHO recommends R21/Matrix-M vaccine for malaria prevention in updated advice on immunization. www.who.int/news/item/02-10-2023-who-recommends-r21-matrix-m-vaccine-for-malaria-prevention-in-updated-advice-on-immunization (2023).
6. March, S. et al. A microscale human liver platform that supports the hepatic stages of *Plasmodium falciparum* and *vivax*. *Cell Host Microbe* **14**, 104–115 (2013).
7. Campo, J. J. et al. RTS,S vaccination is associated with serologic evidence of decreased exposure to *Plasmodium falciparum* liver- and blood-stage parasites. *Mol. Cell Proteomics* **14**, 519–531 (2015).
8. RTS,S Clinical Trials Partnership. Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: final results of a phase 3, individually randomised, controlled trial. *Lancet* **386**, 31–45 (2015).
9. World Health Organization. Monoclonal antibodies for malaria prevention: preferred product characteristics and clinical development considerations. https://www.who.int/publications/i/item/9789240070981 (2023).
10. Daily, J. P. Monoclonal antibodies — a different approach to combat malaria. *N. Engl. J. Med.* **387**, 460–461 (2022).
11. Gaudinski, M. R. et al. A monoclonal antibody for malaria prevention. *N. Engl. J. Med.* **385**, 803–814 (2021).
12. Kayentao, K. et al. Safety and efficacy of a monoclonal antibody against malaria in Mali. *N. Engl. J. Med.* **387**, 1833–1842 (2022).
13. Wu, R. L. et al. Low-dose subcutaneous or intravenous monoclonal antibody to prevent malaria. *N. Engl. J. Med.* **387**, 397–407 (2022).
14. Dame, J. B. et al. Structure of the gene encoding the immunodominant surface antigen on the sporozoite of the human malaria parasite *Plasmodium falciparum*. *Science* **225**, 593–599 (1984).
15. Zavala, F., Cochrane, A. H., Nardin, E. H., Nussenzweig, R. S. & Nussenzweig, V. Circumsporozoite proteins of malaria parasites contain a single immunodominant region with two or more identical epitopes. *J. Exp. Med.* **157**, 1947–1957 (1983).
16. Wang, L. T. et al. A potent anti-malarial human monoclonal antibody targets circumsporozoite protein minor repeats and neutralizes sporozoites in the liver. *Immunity* **53**, 733–744 (2020).
17. Kisalu, N. K. et al. A human monoclonal antibody prevents malaria infection by targeting a new site of vulnerability on the parasite. *Nat. Med.* **24**, 408–416 (2018).
18. Zeeshan, M. et al. Genetic variation in the *Plasmodium falciparum* circumsporozoite protein in India and its relevance to RTS,S malaria vaccine. *PLoS ONE* **7**, e43430 (2012).
19. Weber, J. L. & Hockmeyer, W. T. Structure of the circumsporozoite protein gene in 18 strains of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **15**, 305–316 (1985).
20. Oyen, D. et al. Structural basis for antibody recognition of the NANP repeats in *Plasmodium falciparum* circumsporozoite protein. *Proc. Natl Acad. Sci. USA* **114**, E10438–E10445 (2017).
21. Flores-Garcia, Y. et al. Optimization of an in vivo model to study immunity to *Plasmodium falciparum* pre-erythrocytic stages. *Malar. J.* **18**, 426 (2019).
22. Otsubo, R. & Yasui, T. Monoclonal antibody therapeutics for infectious diseases: beyond normal human immunoglobulin. *Pharmacol. Ther.* **240**, 108233 (2022).
23. Kelley, B., Renshaw, T. & Kamarck, M. Process and operations strategies to enable global access to antibody therapies. *Biotechnol. Prog.* **37**, e3139 (2021).
24. Regules, J. A. et al. Fractional third and fourth dose of RTS,S/AS01 malaria candidate vaccine: a phase 2a controlled human malaria parasite infection and immunogenicity study. *J. Infect. Dis.* **214**, 762–771 (2016).
25. Triller, G. et al. Natural parasite exposure induces protective human anti-malarial antibodies. *Immunity* **47**, 1197–1209 (2017).
26. Tan, J. et al. A public antibody lineage that potently inhibits malaria infection through dual binding to the circumsporozoite protein. *Nat. Med.* **24**, 401–407 (2018).
27. Imkeller, K. et al. Antihomotypic affinity maturation improves human B cell responses against a repetitive epitope. *Science* **360**, 1358–1362 (2018).
28. Murugan, R. et al. Evolution of protective human antibodies against *Plasmodium falciparum* circumsporozoite protein repeat motifs. *Nat. Med.* **26**, 1135–1145 (2020).
29. Cox, R. J. et al. An early humoral immune response in peripheral blood following parenteral inactivated influenza vaccination. *Vaccine* **12**, 993–999 (1994).
30. Radbruch, A. et al. Competence and competition: the challenge of becoming a long-lived plasma cell. *Nat. Rev. Immunol.* **6**, 741–750 (2006).
31. Nutt, S. L., Hodgkin, P. D., Tarlinton, D. M. & Corcoran, L. M. The generation of antibody-secreting plasma cells. *Nat. Rev. Immunol.* **15**, 160–171 (2015).
32. Murugan, R. et al. Clonal selection drives protective memory B cell responses in controlled human malaria infection. *Sci. Immunol.* **3**, eaap8029 (2018).
33. Foquet, L. et al. Vaccine-induced monoclonal antibodies targeting circumsporozoite protein prevent *Plasmodium falciparum* infection. *J. Clin. Invest.* **124**, 140–144 (2014).
34. Flores-Garcia, Y. et al. The *P. falciparum* CSP repeat region contains three distinct epitopes required for protection by antibodies in vivo. *PLoS Pathog.* **17**, e1010042 (2021).
35. Raghunandan, R. et al. Characterization of two in vivo challenge models to measure functional activity of monoclonal antibodies to *Plasmodium falciparum* circumsporozoite protein. *Malar. J.* **19**, 113 (2020).
36. Suscovich, T. J. et al. Mapping functional humoral correlates of protection against malaria challenge following RTS,S/AS01 vaccination. *Sci. Transl. Med.* **12**, eabb4757 (2020).
37. Chatterjee, D. et al. Avid binding by B cells to the *Plasmodium* circumsporozoite protein repeat suppresses responses to protective subdominant epitopes. *Cell Rep.* **35**, 108996 (2021).

38. Livingstone, M. C. et al. In vitro and in vivo inhibition of malaria parasite infection by monoclonal antibodies against *Plasmodium falciparum* circumsporozoite protein (CSP). *Sci. Rep.* **11**, 5318 (2021).
39. Oyen, D. et al. Cryo-EM structure of *P. falciparum* circumsporozoite protein with a vaccine-elicited antibody is stabilized by somatically mutated inter-Fab contacts. *Sci. Adv.* **4**, eaau8529 (2018).
40. Pholcharee, T. et al. Diverse antibody responses to conserved structural motifs in *Plasmodium falciparum* circumsporozoite protein. *J. Mol. Biol.* **432**, 1048–1063 (2020).
41. Pholcharee, T. et al. Structural and biophysical correlation of anti-NANP antibodies with in vivo protection against *P. falciparum*. *Nat. Commun.* **12**, 1063 (2021).
42. Martin, G. M. et al. Affinity-matured homotypic interactions induce spectrum of PfCSP structures that influence protection from malaria infection. *Nat. Commun.* **14**, 4546 (2023).
43. Dennison, S. M. et al. Breadth of human monoclonal antibodies isolated from RTS,S/AS01 vaccinees binding to *Plasmodium falciparum* circumsporozoite protein antigens. *Biophys. J.* **116**, 480a (2019).
44. Thai, E. et al. Molecular determinants of cross-reactivity and potency by VH3-33 antibodies against the *Plasmodium falciparum* circumsporozoite protein. *Cell Rep.* https://doi.org/10.1016/j.celrep.2023.113330 (2023).
45. Wang, L. T. et al. Protective effects of combining monoclonal antibodies and vaccines against the *Plasmodium falciparum* circumsporozoite protein. *PLoS Pathog.* **17**, e1010133 (2021).
46. Langowski, M. D. et al. Restricted valency (NPNA)n repeats and junctional epitope-based circumsporozoite protein vaccines against *Plasmodium falciparum*. *NPJ Vaccines* **7**, 13 (2022).
47. McNamara, H. A. et al. Antibody feedback limits the expansion of B cell responses to malaria vaccination but drives diversification of the humoral response. *Cell Host Microbe* **28**, e7 (2020).
48. Vijayan, K. et al. Antibody interference by a non-neutralizing antibody abrogates humoral protection against *Plasmodium yoelii* liver stage. *Cell Rep.* **36**, 109489 (2021).
49. Vijay, R. et al. Infection-induced plasmablasts are a nutrient sink that impairs humoral immunity to malaria. *Nat. Immunol.* **21**, 790–801 (2020).
50. Kucharska, I. et al. High-density binding to *Plasmodium falciparum* circumsporozoite protein repeats by inhibitory antibody elicited in mouse with human immunoglobulin repertoire. *PLoS Pathog.* **18**, e1010999 (2022).
51. Xu, Y. et al. Structure, heterogeneity and developability assessment of therapeutic antibodies. *mAbs* **11**, 239–264 (2019).
52. Fernández-Quintero, M. L. et al. Assessing developability early in the discovery process for novel biologics. *mAbs* **15**, 2171248 (2023).
53. Gunasekaran, K., Hagler, A. T. & Gierasch, L. M. Sequence and structural analysis of cellular retinoic acid-binding proteins reveals a network of conserved hydrophobic interactions. *Proteins* **54**, 179–194 (2004).
54. Kerwin, B. A. et al. Framework mutations of the 10-1074 bnAb increase conformational stability, manufacturability, and stability while preserving full neutralization activity. *J. Pharm. Sci.* **109**, 233–246 (2020).
55. Zalevsky, J. et al. Enhanced antibody half-life improves in vivo activity. *Nat. Biotechnol.* **28**, 157–159 (2010).
56. Kurtovic, L. et al. Multifunctional antibodies are induced by the RTS,S malaria vaccine and associated with protection in a phase 1/2a trial. *J. Infect. Dis.* **224**, 1128–1138 (2021).
57. White, M. T. et al. The relationship between RTS,S vaccine-induced antibodies, $CD4^+$ T cell responses and protection against *Plasmodium falciparum* infection. *PLoS ONE* **8**, e61395 (2013).
58. Kester, K. E. et al. A phase I/IIa safety, immunogenicity, and efficacy bridging randomized study of a two-dose regimen of liquid and lyophilized formulations of the candidate malaria vaccine RTS,S/AS02A in malaria-naïve adults. *Vaccine* **25**, 5359–5366 (2007).
59. Ockenhouse, C. F. et al. Ad35.CS.01 - RTS,S/AS01 heterologous prime boost vaccine efficacy against sporozoite challenge in healthy malaria-naïve adults. *PLoS ONE* **10**, e0131571 (2015).
60. Chaudhury, S. et al. The biological function of antibodies induced by the RTS,S/AS01 malaria vaccine candidate is determined by their fine specificity. *Malar. J.* **15**, 301 (2016).
61. Pallikkuth, S. et al. A delayed fractionated dose RTS,S AS01 vaccine regimen mediates protection via improved T follicular helper and B cell responses. *eLife* **9**, e51889 (2020).
62. Langowski, M. D. et al. Optimization of a *Plasmodium falciparum* circumsporozoite protein repeat vaccine using the tobacco mosaic virus platform. *Proc. Natl Acad. Sci. USA* **117**, 3114–3122 (2020).
63. Chatterjee, D. & Cockburn, I. A. The challenges of a circumsporozoite protein-based malaria vaccine. *Expert Rev. Vaccines* **20**, 113–125 (2021).
64. Cockburn, I. A. & Seder, R. A. Malaria prevention: from immunological concepts to effective vaccines and protective antibodies. *Nat. Immunol.* **19**, 1199–1211 (2018).
65. Wang, L. T. et al. The light chain of the L9 antibody is critical for binding circumsporo-zoite protein minor repeats and preventing malaria. *Cell Rep.* **38**, 110367 (2022).
66. Calvo-Calle, J. M., Mitchell, R., Altszuler, R., Othoro, C. & Nardin, E. Identification of a neutralizing epitope within minor repeat region of *Plasmodium falciparum* CS protein. *NPJ Vaccines* **6**, 10 (2021).
67. Hou, N. et al. Low-complexity repetitive epitopes of *Plasmodium falciparum* are decoys for humoural immune responses. *Front. Immunol.* **11**, 610 (2020).
68. Rénia, L. & Goh, Y. S. Malaria parasites: the great escape. *Front. Immunol.* **7**, 463 (2016).
69. Schofield, L. The circumsporozoite protein of *Plasmodium*: a mechanism of immune evasion by the malaria parasite? *Bull. World Health Organ.* **68**, 66–73 (1990).
70. Schofield, L. & Uadia, P. Lack of Ir gene control in the immune response to malaria. I. A thymus-independent antibody response to the repetitive surface protein of sporozo-ites. *J. Immunol.* **144**, 2781–2788 (1990).
71. Raghavan, M. et al. Antibodies to repeat-containing antigens in *Plasmodium falciparum* are exposure-dependent and short-lived in children in natural malaria infections. *eLife* **12**, e81401 (2023).
72. Friedman-Klabanoff, D. J. et al. Epitope-specific antibody responses to a *Plasmodium falciparum* subunit vaccine target in a malaria-endemic population. *J. Infect. Dis.* **223**, 1943–1947 (2021).
73. Anders, R. F. et al. Antigenic repeat structures in proteins of Plasmodium falciparum. *Ciba Found. Symp.* **119**, 164–183 (1986).
74. Dennison, S. M. et al. Qualified biolayer interferometry avidity measurements distinguish the heterogeneity of antibody interactions with *Plasmodium falciparum* circumsporozoite protein antigens. *J. Immunol.* **201**, 1315–1326 (2018).
75. Kucharska, I. et al. Structural ordering of the *Plasmodium berghei* circumsporozoite protein repeats by inhibitory antibody 3D11. *eLife* **9**, e59018 (2020).
76. RTS,S/AS01E malaria vaccine (MoSQUIRIX*) Children living in malaria-endemic re-gions: little efficacy, poorly documented harms. *Prescrire Int.* **26**, 5–8 (2017).
77. Fochesato, M., Dendouga, N. & Boxus, M. Comparative preclinical evaluation of AS01 versus other Adjuvant Systems in a candidate herpes zoster glycoprotein E subunit vaccine. *Hum. Vaccin. Immunother.* **12**, 2092–2095 (2016).

[1]Atreca, Inc., San Carlos, CA, USA. [2]Department of Molecular Microbiology and Immunology, Malaria Research Institute, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA. [3]Initium Therapeutics, Inc., Natick, MA, USA. [4]Just – Evotec Biologics, Seattle, WA, USA. [5]Duke Center for Human Systems Immunology, Department of Surgery, Duke University, Durham, NC, USA. [6]BioNTech US, Inc., Cambridge, MA, USA. [7]PATH Center for Vaccine Innovation and Access, Washington DC, USA. [8]Walker Bioscience, Carlsbad, CA, USA. [9]GSK, Rixensart, Belgium. [10]Centivax, Inc., South San Francisco, CA, USA. [11]Center for Enabling Capabilities, Walter Reed Army Institute of Research, Silver Spring, MD, USA. [12]Walter Reed Army Institute of Research, Silver Spring, MD, USA. [13]Nuevocor Pte. Ltd, Singapore, Singapore. [14]Paramune, Inc., San Carlos, CA, USA. [15]Division of Immunology and Rheumatology, Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA. [16]Departments of Immunology, Molecular Genetics and Microbiology, Human Vaccine Institute, Duke University, Durham, NC, USA. ✉e-mail: kwilliams@atreca.com; danielemerling@biosimplify.com

Methods

All research complied with appropriate regulations, committees and institutional requirements as specifically noted in the relevant subsections below.

Vaccinees, PB isolation and IgG sequencing

PBMCs were collected as part of a phase 2a clinical trial evaluating the RTS,S/A01 vaccine with fractional third and fourth doses[24]. The protocol was approved by the Walter Reed Army Institute of Research Institutional Review Board and the Western Institutional Review Board, and written informed consent was obtained from each subject before study procedures were initiated (ClinicalTrials.gov identifier: NCT01857869). PBMC samples obtained from trial participants or this study were used exhaustively and are not available.

PB isolation, cloning and sequencing were performed using BD FACSDiva software v.8.0.3 and the previously published protocol[78] with the following modifications. PBMCs were stained with the following mAbs and dilutions: anti-CD3-FITC (BioLegend, catalog no. 300440, clone UCHT1, 1:100), anti-CD14-FITC (BioLegend, catalog no. 325604, clone HCD14, 1:100), anti-CD19-BV421 (BioLegend, catalog no. 302234, clone HIB19, 1:100), anti-CD20-PE/cy7 (Biolegend, catalog no. 302312, clone 2H7, 1:100), anti-CD27-BV510 (BioLegend, catalog no. 302836, clone O323, 1:50), anti-CD38-A647 (BioLegend, catalog no. 303514, clone HIT2, 1:100), anti-IgA-FITC (Miltenyi, catalog no. 130- 113-475, clone IS11-8E10, 1:50), anti-IgM-FITC (BioLegend, catalog no. 314506, clone MHM-88, 1:50) and anti-IgD-FITC (Biolegend, catalog no. 348206, clone IA6-2, 1:50). IgG^+ PBs were single-cell sorted into 96-well polymerase chain reaction (PCR) plates containing hypotonic buffer (330 nM dNTPs (NEB, catalog no. N0447L), 1 µg ml^{-1} bovine serum albumin (NEB, catalog no. B9000S), 2 mM DTT (Sigma-Aldrich, catalog no. 43816), 0.5% IGEPAL-630 (Sigma-Aldrich, catalog no. I8896) and 500 units ml^{-1} Ribolock (Thermo Fisher Scientific, catalog no. EO0384)) based on gating for $CD3^-CD14^-CD19^+CD20^-CD27^+CD38^+ + IgA^-IgM^-IgD^-$ cells (Supplementary Fig. 9). Sequencing of IgG messenger RNA isolated from single-cell sorted PBs was performed with the following modifications: desthiobiotinylated Oligo(dT) and Maxima H-Reverse Transcriptase (Thermo Fisher Scientific, catalog no. EP0753) were used for reverse transcription (RT), complementary DNA was extracted using Dynabeads MyOne C1 Streptavidin beads (Thermo Fisher Scientific, catalog no. 65001), concentrations of final NGS library preparations were determined using quantitative PCR (qPCR; KAPA SYBR FAST qPCR Kit for Titanium, Kapa Biosystems) and natively paired IgG heavy- and light-chain amplicons were sequenced using Roche GS FLX + 454 Titanium sequencing with system Software v.2.9.

DNA barcode assignment and sequence assembly were performed as previously described[78] except for the following modification: a minimum coverage of ten reads was required for each heavy- and light-chain assembly. Both heavy- and light-chain reads were required to assemble unique contigs within a well. In cases where there was one or more contig, we rejected that well unless one of the contigs constituted 90% of reads.

Sequence, lineage and repertoire feature analyses

Germline assignments and determination of SHM levels. Variable (V), diversity (D) and joining (J) gene segment assignment and mutation identification were performed using an implementation of somatic diversification analysis[79] and the IMGT human immunoglobulin germline database release IMGT_202031 (ref. 80). Nucleotide SHM substitutions were counted by aligning the heavy and light variable domains (start of framework 1 to end of framework 4) with a hidden Markov model that includes states for germline aligning regions (VDJ for heavy, VJ for light) and N nucleotide regions, and that counts substitutions with respect to the germline sequence in just the aligned portion (not including the rare, observed indels). IgG isotype (IgG1–IgG4) assignment was performed by alignment of sequence 3' of framework 4 to the IMGT human Ig constant-region sequences from IMGT_202031 (ref. 80).

CDR3 and lineage assignments. Complementarity-determining region 3 (CDR3) sequences were defined by the Kabat annotation plus the first amino acid residue of framework 4, from which CDR3 lengths were calculated. Natively paired IgG sequence clones were assigned to the same lineage if derived from the same vaccinee, had the same IGHV and IGK/LV germline gene assignments, the same heavy- and light-chain CDR3 (H3 and L3, respectively) length and at least 75% nucleic acid sequence identity across concatenated H3 and L3. In some cases, clones with IGHV3-33 and IGHV3-30 (germline genes with high sequence identity) met all the criteria for assignment to the same lineage, except for the difference in IGHV. In these cases clones were assigned to the same lineage. Lineages were assigned a rank-size based on lineage frequency (number of PBs expressing clones in the lineage divided by total number of PBs in the repertoire). Lineages with the same number of PBs were reported to have the same rank-size.

Convergence, clonality and recall. Two IgG clones were defined as convergent if derived from different vaccinees, had the same IGHV and IGK/LV germline gene assignments, the same heavy- and light-chain CDR3 (H3 and L3, respectively) length and at least 85% BLOSUM62-weighted amino acid sequence similarity between concatenated H3 and L3. A lineage was defined as convergent with another lineage if they were derived from different vaccinees and if there was at least one IgG clone in the first lineage that was convergent with at least one in the second. Clonality was summarized as normalized entropy across all lineages in each P3D vaccinee repertoire. Specifically, the sum over i in 1 to N of $(Ki/N \times \log(Ki/N))/\log(N)$, where N equals the number of lineages in the repertoire and Ki is the size of each lineage as the number of PBs. Normalized entropy takes values between 0 and 1 inclusive, where 0 implies that a single lineage is totally dominant in abundance and 1 implies that some sets of lineages $N > 1$ are all equally abundant. Recalled lineages were defined as those with at least one PB antibody clone observed in both P3D and P4D repertoires among vaccinees ($n = 17$) from whom at least 100 PBs were sequenced from P4D PBMC samples.

Selection of lineages and clones for CSP ELISA screening library

Plasmablast lineages ($n = 369$) were selected for initial screening of CSP reactivity as described above (Supplementary Fig. 8a and Supplementary Table 2). A specific clone from each lineage was chosen for recombinant expression and screening based on one or more of the following properties: (1) the clone was the most frequently observed across the PBs of the lineage ('dominant clone'); (2) the antibody sequence is convergent with at least one other sequence where convergence is described in 'Convergence, clonality and recall'; (3) the clone was selected from the 'leafiest descent' clade of a phylogenetic tree, where each terminal clade is rank ordered by size according to the number of leaves (that is, sequences); and (4) the clone has the greatest number of nucleic acid mutations from germline among all clones in the lineage ('most mutated clone'). Three clones (0.8% of the screening library) did not meet any of these criteria due to errors that were not detected until after screening had occurred. Supplementary Table 5 shows which criteria apply to each selected clone, with some noted as 'tie'. In this case, two clones from the same lineage met one of the criteria above and we made a random selection between the two. Proportions of antibodies that meet each criterion from protected and unprotected vaccinees are not statistically different to those seen for all antibodies in the screening library, Fisher's exact test (Supplementary Table 6).

Recombinant antibody production

Antibody syntheses were performed by LakePharma, Inc. Each gene sequence was cloned into LakePharma's proprietary high-expression

mammalian vector. Variable regions sequences were synthesized and subcloned into expression vectors containing human heavy-chain IgG1 and appropriate human kappa or lambda light-chain constant-region sequences. Each completed construct was sequence confirmed before plasmid production scale-up. Suspension HEK293 cells (ATCC) were seeded in a shake flask and expanded using serum-free, chemically defined medium. On the day of transfection the expanded cells were seeded into a new flask with fresh medium. Each DNA construct was transiently transfected into HEK293 cells using a cationic lipid transfection method. Cells were maintained as a batch-fed culture until the end of the production run. Conditioned medium from the transient production run was harvested and clarified by centrifugation and filtration. The supernatant was loaded over a Protein A column pre-equilibrated with binding buffer. Washing buffer was passed through the column until the optical density (OD_{280}) value (NanoDrop, Thermo Scientific) was measured as zero. The target protein was eluted with a low-pH buffer, fractions were collected and the OD_{280} value of each fraction was recorded. Fractions containing the target protein were pooled and filtered through a 0.2-µm membrane filter. Purified antibodies were dialyzed against PBS and analyzed with LabChip GXII. Endotoxin measurements were performed using the chromogenic limulus amebocyte lysate method with Pyrochrome (Associates of Cape Cod).

CSP, NANP peptide and C-terminal peptide ELISA

ELISA methodologies have previously been described[24]. mAbs were mapped to CSP using a nearly-full-length CSP, the NANP6 peptide and a CSP C-terminal peptide (Pf16)[81]. All antibodies were evaluated at either 0.15 or 0.04 µg ml^{-1} concentrations and classified as 'positive', 'negative' or 'indeterminant'. OD values were determined using Biotek Gen5 software v.2. ELISA OD was converted to fold-induction over the average of four negative control antibodies run in each experiment. The range of OD responses observed in each experiment was then used to determine a borderline indeterminant range for that experiment. MAbs were considered negative if the OD was less than the average negative control mAb OD + 3× standard deviations. MAbs with ODs that exceeded the 'negative' threshold but were within 20% of the OD range for the experiment were classified as indeterminant. Any antibody OD above the experimental indeterminant threshold was classified as positive.

HBsAg ELISA

Antibodies and the positive-control PC3 (starting at 150 pM) were evaluated for HBsAg reactivity using a four-point, 1:3 dilution series prepared in duplicate and tested using the MONOLISA Anti-HBs EIA kit (Bio-Rad, catalog no. 25220). The maximal stock input was 10% of the purified total volume for each of the 139 tested antibodies. If the antibody required <10% of the purified total volume the starting concentration was individually adjusted to 300 nM; otherwise the starting concentration was based on the protein amount included in 10% of the purified total volume. The cutoff calibrator from the kit was performed in quadruplicate while both negative controls were each performed in duplicate. Antibodies were considered HBsAg positive if the signal met the cutoff calibrator criteria for at least one concentration ≤30 nM. Antibodies were considered 'borderline' HBsAg reactive if the signal was negative at concentrations tested ≤30 nM but did meet the cutoff calibrator criteria for any concentration >30 nM. Antibodies were considered negative if the signal did not meet the cutoff calibrator criteria for any concentration tested.

Selection of mAbs for initial characterization in the mouse sporozoite-challenge model

Among the 102 mAbs that were reactive in the NANP6 peptide ELISA (Fig. 1d), 69 were selected to be screened in vivo based on IGHV representation, vaccine protection status and SHM levels. These 102 mAbs were divided into 11 groups based on their different IGHV (IGHV1–2, 1–69, 1–8, 3–15, 3–23, 3–30, 3–33, 3–48, 3–49, 3–7 and 5–51). We selected at least half of the mAbs from each IGHV group; these mAbs originated from 26 protected and eight unprotected. vaccinees. Only mAbs with high levels of SHM (≥20 nucleotide mutations from germline per antibody) were selected, except for the sets that contained IGHV3-33 and 3-49, from which some mAbs with low levels of SHM (<20 nucleotide mutations from germline per antibody) were also included. Two mAbs did not express sufficient material for testing in vivo.

Among the 20 mAbs that were reactive in the C-terminal (Pf16) ELISA (Fig. 1d), 11 of the 12 originated from protected vaccinees and were screened in vivo. These mAbs included all of the IGHV germline genes observed among C-terminal (Pf16) binders from protected vaccinees (IGHV3–11, 3–21, 3–30, 3–48 and 4–59). One mAb did not express sufficient material for testing in vivo.

Selection of mAbs characterized in SPR binding analyses

All of the mAbs that originated from protected vaccinees ($n = 36$) and that demonstrated ≥90% inhibition in the initial sporozoite liver burden mouse model screen were selected for further binding analyses, except for one that was observed to be reactive in the HBsAg ELISA (AB-000239; Supplementary Table 2). For each of these 35 mAbs (representing 35 unique lineages) the original hit mAb was selected if it contained no high-risk liability (that is, odd number of cysteines in CDRs, any canonical N-linked glycosylation sites in CDRs, Fv net charge >9 (at pH 5.5) or hydrophobicity index >6.5). In cases where the original hit had one or more high-risk liabilities, a different clone was selected from the lineage. More than one clone was selected from lineages with extensive interclonal sequence diversity; this was done using the following algorithm: (1) query each clone of lineage in leafiest descent order; (2) skip any clones with more than one high-risk liability; (3) skip any clones overly similar in amino acid sequence to clones already picked and where the distance between clones was determined as the fraction of distinct CDR amino acid positions using the BLOSUM62 matrix (≤0); and (4) adjust the distance that is acceptable between clones so that a total of 141 clones were ultimately selected from the 35 lineages (Supplementary Table 4).

High-throughput SPR

The binding kinetics measurements of mAb interaction with CSP antigens were made using the Carterra LSA high-throughput SPR platform with Epitope Software v.1.5 and CMD200M sensor chips (Carterra) at 25 °C. The antigen panel included recombinant CSP and synthetic peptides NPNA3 (NPNANPNANPNA) and NANP6 (NANPNANPNANPNANPNANPNANP), junctional peptide (KQPADGNPDPNANPN), NPDPNANP2NVDP (NPDPNANPNVDPNANP) and NVDP3NANP2 (NVDPNANPNVDPNANPNVDP), which were custom made by CPC Scientific. Except for NANP6, all other peptides were acetylated at N termini and amidated at C termini. NANP6 contained an N-terminal biotin-aminohexanoic acid tag and an unmodified C terminus. Two microfluidic modules, a 96-channel print-head (96PH) and a single flow cell (SFC) were used to deliver liquids onto the sensor chip. A single analyte antigen was titrated in each assay against the immobilized antibodies.

Immobilization of antibodies onto CMD200M chips depended on the type of analyte used during titration. In assays involving recombinant CSP used as an analyte, a goat anti-human IgG Fc antibody (Millipore, Sigma-Aldrich AP113, lot no. 3027077) was first immobilized onto the chip through amine coupling. The chip was then activated by 100 mM *N*-hydroxysuccinimide and 400 mM 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) (GE Healthcare, mixed 1:1:1 with 0.1 M MES buffer pH 5.5) for 600 s and followed by immobilization of anti-human IgG Fc (in 10 mM sodium acetate pH 4.5) at 50 µg ml^{-1} for 900 s. Unreactive esters were quenched with a 600-s injection of 1 M ethanolamine-HCl pH 8.5. The chip was then exposed to double pulses (30 s per pulse) of 10 mM glycine pH 2.0.

CSP-specific mAbs were then captured on anti-Hu IgG Fc surfaces by injection of mAbs at either 10 or 5 μg ml^{-1} for 600 s using the 96PH, with 1× HBSTE buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA and 0.01% Tween-20) as running buffer and antibody diluent. When CSP-peptide antigens were used as analytes, the chip was activated by *N*-hydroxysuccinimide/EDC for 600 s followed by direct immobilization of CSP-specific mAbs (in 10 mM sodium acetate pH 4.5) injected at either 10 or 5 μg ml^{-1} for 600 s using the 96PH. Unreactive esters were quenched with a 600-s injection of 1 M ethanolamine-HCl pH 8.5, then 45 cycles of 1× HBSTE buffer injections with 1× HBSTE as running buffer were used to wash off nonspecifically bound IgG overnight from the sensor chip surface without the use of regeneration buffer. Except for the capture of mAbs by anti-human IgG Fc and washing of nonspecifically bound IgG, the running buffer was 10 mM MES buffer pH 5.5 with 0.01% Tween-20. Unless specified above, the steps were done using the SFC.

During the initial screening each mAb dilution was immobilized onto two separate spots of the same chip, enabling duplicate measurements. The engineered variants were immobilized onto three different spots enabling triplicate measurements.

A twofold dilution series of the antigen was prepared in 1× HBSTE buffer. The top concentration for full-length CSP and all CSP-peptide antigens was 8 μg ml^{-1} (0.25 μM for CSP, 2.92 μM for NANP6, 6.41 μM for NPNA3, 3.76 μM for NVDP3NANP2, 4.70 μM for NPDPNANPNVDPNANP and 5.03 μM for N-interface). The antigen at different concentrations was then injected using SFC onto the chip surface, from the lowest to the highest concentration without regeneration, including eight injections of buffer before the lowest nonzero concentration for signal stabilization. For each concentration, data collection involved 120 s of baseline step and 900 s of dissociation steps. The threshold for measurable off-rates was 10^{-5} s^{-1}. Due to the sensitivity of the instrument, off-rates slower than 10^{-5} s^{-1} were associated with large standard errors and could not be confidently determined. The duration of the association step was 240 s for full-length CSP and NANP6 antigens and 300 s for all other CSP-peptide antigens. For all assays the running buffer for titration was 1× HBSTE.

The kinetics titration data collected were first preprocessed in NextGenKIT (Carterra) software, including reference subtraction, buffer subtraction and data smoothing. The data were then exported and analyzed using the TitrationAnalysis tool (https://zenodo.org/record/7998652) developed in-house[82] (https://gatesopenresearch.org/articles/7-107/v1). The specific binding time courses for each antibody construct immobilized on different spots were fitted to a 1:1 Langmuir model to derive k_a ('k_{on}', association rate constant), k_d ('k_{off}', dissociation rate constant) and K_d (dissociation constant) values. The K_d values determined for antigens with multiple epitope repeats include the avidity effect. The average values of duplicate measurements were reported for each antibody–antigen pair from the initial screening panel. For the engineered variants, average values of triplicate measurements were reported with the following data acceptance criteria: (1) standard error of the estimated k_{on}, k_{off} and K_d in each replicate ≤20% and (2) fold change for all three parameters within the triplicate ≤3.

In vivo functional assessments

Mouse studies used 6–8-week-old C57BL/6 female mice (Charles River Labs), maintained at the animal facility of the Johns Hopkins Bloomberg School of Public Health. Mouse housing was maintained at 40–60% relative humidity and a temperature of 68–79 °F), with at least ten room air changes per hour and a 14/10-h light/dark cycle. No animals or data points were excluded from analyses. Experiments were performed in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The protocol was approved by the Animal Care and Use Committee of Johns Hopkins University (protocol nos. MO18H419 and MO21H417).

Sporozoite-challenge liver burden mouse models. In the first functional screen (data are shown in Fig. 2a and Supplementary Table 2), mice (n = 10) were administered via intravenous tail vein either 100 or 300 μg of mAb (dose specified in Supplementary Table 2) and were challenged 16 h later via intravenous tail vein with 2,000 *P. berghei* transgenic sporozoites expressing the full *P. falciparum* CSP. Forty-two hours later, mice were euthanized and their livers excised to extract RNA to perform qPCR with RT to measure plasmodial 18S ribosomal RNA using forward primer 5′-TGGGAGATTGGTTTTGACGTTTATGT-3′ and reverse primer 5′-AAGCATTAAATAAAGCGAATACATCCTTAC-3′ as previously described[83]. Parasite loads were expressed as *P. berghei* 18S rRNA copy number, and percentage inhibition of load was calculated in comparison with negative controls.

All other liver burden assays were performed as previously described[21]. *Anopheles stephensi* mosquitoes infected with transgenic *P. berghei* sporozoites expressing *P. falciparum* CSP and luciferase were maintained in an incubator at 19 °C. Sporozoites from mosquitoes were harvested at days 20–23 post infection in 2% HBSS-FBS. Mice (n = 5) were intravenously (tail vein) administered 100 μg of antibody, unless otherwise indicated, and challenged 16 h later with 2,000 sporozoites injected via the tail vein. Control mice received either irrelevant or no antibodies. Forty-two hours following challenge, mice were injected with 100 μl of D-luciferin (30 mg ml^{-1}), anesthetized with isoflurane and bioluminescence expressed by liver parasites was measured using an IVIS Spectrum imager with Living Image v.3.2 software (Perkin Elmer). Results are expressed as photons s^{-1}.

Mosquito-bite challenge parasitemia mouse models. The mosquito-bite challenge for evaluation of sterile protection was performed as described by Flores-Garcia et al.[21]. Mice (n = 10) were passively immunized with 150 μg of antibody, anesthetized 16 h later and placed for 10 min on the top of cages containing five mosquitoes infected with *P. berghei* sporozoites expressing *P. falciparum* CSP and luciferase. From days 4–10 following challenge, blood smears stained with Giemsa were observed under a light microscope to determine the appearance of parasitemia. Control mice receiving irrelevant or no antibodies were challenged similarly.

Assessment of mAb concentrations in serum samples

Blood was collected by retro-orbital bleed 15 h following antibody administration (1 h before infection). Capture antibody (AffiniPure mouse anti-human IgG Fc fragment specific, Jackson ImmunoResearch, no. 209-005-098) was adsorbed overnight at 21 °C onto 96-well polystyrene microplates (Immuno Plate Maxisorp, Thermo Fisher Scientific, no. 439454) in PBS (Dulbecco's PBS, without calcium and magnesium, sterile, pH 7.4; Wisent, no. 311-425-LL) and then washed three times in wash buffer (0.05% Tween 20 (Sigma, no. P2287) in PBS). Microplates were blocked for 1 h at 21 °C with assay buffer (1% bovine serum albumin (Blocker BSA, ThermoFisher, no. 37525) in wash buffer). After washing with wash buffer three times, serum samples from mice and control standards were added in duplicate at serial dilutions in normal mouse serum and then further diluted 100-fold in assay buffer before incubation for 1 h at 21 °C. Control standards consisted of AB-000317 serially diluted at 1.6-fold increments. from 0.146 to 25.6 μg ml^{-1}. Microplates were then washed three times with wash buffer and incubated with mouse monoclonal anti-human IgG antibody conjugated to horse radish peroxidase-conjugated clone JDC-101 (Southern Biotech, no. 9040-05) in assay buffer for 1 h at 21 °C. Following three washes with wash buffer, peroxidase substrate (TMB, Bio-Rad, no. 1721068) was added followed by a stop solution (650 nm of TMB stop solution, Southern Biotech, no. 0413-01L). Absorbance was measured at 650 nm (Molecular Devices microplate reader using SoftMaxPro GxP v.6.5.1), and concentrations of human IgG in test samples were calculated using the standard curve generated from the control antibody by interpolation of OD values on the five-parameter logistic standard curve (derived from

mean ODs of duplicate standard samples) and adjusted according to their corresponding dilution factor. Final sample concentrations were then determined by calculating the average of all concentrations for a sample obtained within the range of the standard curve.

Developability characterization assays

Sample preparation. Samples were first buffer exchanged against ten diavolumes of 20 mM sodium phosphate and 150 mM sodium chloride pH 7.1 (PBS) using a centrifugal filter with a 30-kDa molecular weight cutoff (Amicon), then normalized to 1 mg ml^{-1} using a Lunatic protein concentration plate format instrument (Unchained Labs).

Biophysical analysis. DSF. DSF was used to determine thermal transition temperature(s) and weighted shoulder scores according to a method previously described[54].

Thermal hold. Thermal hold was assessed by placing 30 µl of sample in a PCR plate (Bio-Rad) in replicate wells, one plate column for each sample. The plate was heated using a C1000 thermocycler (Bio-Rad) set to apply gradient heat from 69 to 74 °C for 5 min. Samples were transferred to a clear 384-well plate (Fisher Scientific Company) and analyzed for turbidity by reading absorbance at 350 nm using a Spectrostar Nano plate reader (BMG Labtech).

SINS. Self-interaction nanoparticle spectroscopy (SINS) measurements were performed as previously described[84]. Gold nanoparticles (Ted Pella) were conjugated overnight with an 80:20 ratio of anti-human and anti-goat antibodies (Jackson Immuno Research). Unreacted sites were blocked using an aqueous 0.1% (w/v) polysorbate 20 solution. Conjugated gold nanoparticles were concentrated by centrifugation and 95% of the supernatant was removed. Analysis was carried out in PBS (20 mM phosphate, 150 mM NaCl, pH 7.1) at a protein concentration of 0.05 mg ml^{-1} reacted with 5 µl of concentrated conjugated gold nanoparticles. After a 2-h incubation, absorbance spectra from 400 to 600 nm were collected using a Spectrostar Nano plate reader (BMG Labtech) at 2-nm resolution. The wavelength maximum of the spectrum peak is reported.

Stand-up monolayer absorption chromatography. Stand-up monolayer absorption chromatography measurements were performed as previously described[85]. Retention times were determined using a Dionex UPLC equipped with a Zenix column (Sepax Technologies) and a running buffer comprising 150 mM sodium phosphate pH 7.0.

Low-pH stability. Stability during a low-pH hold was determined as previously described[54]. The increase in high molecular weight of the low-pH-exposed sample versus the control sample is reported.

Relative solubility. Solubility was assessed as previously described[54]. Analysis was done in PBS buffer (20 mM sodium phosphate and 150 mM sodium chloride pH 7.1) and a final polyethylene glycol (PEG) 10,000 concentration ranging from 7.2% to 10.4%. The remaining soluble protein following PEG incubation is reported.

Chemical unfolding. The chemical unfolding assay was completed as previously described[54], with some modifications. After a 3-day incubation in 32 guanidine hydrochloride concentrations, samples were measured on a Fluorescence Innovations SUPR-UV plate reader (excitation, 275 nm; emission, 300–440 nm). The measured fluorescence intensity at 362 nm was corrected for scattering and stray light, the unfolding curve was generated by graphing each corrected intensity against the guanidine hydrochloride concentration and the inflection point was reported.

Variant engineering for clinical candidate and backup

The antibody variable domains of AB-000224, AB-007088 and siblings of both antibodies were analyzed using the Just–Evotec Biologics' Abacus platform. Sequences were placed in a structure-based numbering system (ASN; Supplementary Tables 7 and 8) derived from the AHo[86] system to add structural positional data to the sequence space and produce consistent alignments across diverse sets of antibody variable-domain sequences. Germline background was evaluated to predict expression problems and evaluate potential solutions for sequence-based liabilities. Signal peptides were evaluated based on germline background to improve expression. Sequence-level evaluation was performed to identify potential N-linked glycosylation sites, nontypical cysteine residues, isomerization and deamidation sites that could impact function, covariance violations (CVVs)[53] that could impact stability and tryptophan residues in CDR3s that could lead to degradation. CVV sites were evaluated based on severity, sequence location and whether they would be modified across the variant panel or combinatorially. An empirical multiattribute method (MAM)[87] was run to determine actual posttranslational modification levels to bolster sequence evaluation. The published structure of AB-000224 (6WFY)[41] and a structural model of AB-007088 (built using the Molecular Operating Environment, 2022.02 Chemical Computing Group ULC) were evaluated for surface properties that could impact developability such as hydrophobic patches leading to aggregation, electropositive patches leading to increased serum clearance rates and electronegative patches leading to increased high-concentration viscosity. These structures were also used to evaluate the possible mutation sites for property optimization to determine their potential impact on CSP binding.

For AB-000224, the IGLV2-8 signal peptide was chosen for the lambda light chain and the IGKV1-39 signal peptide for the heavy chain. MAM analysis found one low-level deamidation site that appeared to be in contact with CSP based on CSP-binding complex structures; however, given its low-level occurrence and the fact that a stressed sample did not demonstrate increased deamidation, it was removed from consideration in variant designs. The computed CVV sites were found to be strong across both chains, with the majority being in the heavy chain. AB-000224 siblings AB-007110, AB-007111 and AB-007112, which were also active in vivo, were evaluated at the sequence and structure level to determine any potential beneficial modifications to AB-000224. Positions that differed from AB-000224 were considered optimization sites if they were consistent amongst the siblings or of energetic structural interest. This analysis identified three sites, including the deamidation site, and these three positions were mutated and combined in a single variant. Surface patch analysis found nothing of significance. This analysis resulted in 17 combinatorial variant designs, as depicted in Fig. 6i.

For AB-007088, the IGKV1-39 signal peptide was selected for both the kappa light and heavy chains. MAM analysis found two low-level deamidation sites, although neither appeared structurally able to impact CSP binding given the known CSP-binding complex structures (PDB IDs 5BK0 (ref. 25), 6AXL ref. 20, 6AZM ref. 25, 6BQB ref. 26, 6D01 (ref. 27), 6D0X ref. 27 and 6D11 (ref. 27)), and were not included in variant designs. The computed CVV sites were found to be strong only within the heavy chain. Surface patch analysis found some low levels of minor hydrophobic character, which might have increased hydrophobic interactions but, because their repair would probably have degraded function, they were not included in variant designs. This analysis resulted in five combinatorial variant designs, as depicted in Extended Data Fig. 5i.

Generation and assessment of stable cell lines for production

Sequences were codon optimized, synthesized at ATUM and inserted into transposon expression vector V52 (ref. 88). Constant regions consisted of IGHG1*01 and IGLC7*01 for AB-000224 variants, or IGKC*01 for AB-007088 variants. Each variant was transfected into 20×10^6 CHO-K1 GS KO cells (Horizon Discovery) and maintained in a serum-free, chemically defined CD OptiCHO commercial growth

medium (Gibco) with 4 mM L-glutamine. Transposon DNA (18.75 µg) was combined with 6.25 µg of transposase RNA and mixed with EX-CELL 325 PF CHO (Sigma) to a final volume of 50 µl. Cells were electroporated using the long-duration method[89] in 4-mm gap cuvettes. Cells were electroporated at 200-V voltage, 725-µF resistance and 3,175-Ω capacitance in an Electro Cell Manipulator 630 coupled to a 630B Safety Stand (BTX). Following electroporation, cells were resuspended in 15 ml of fresh growth medium with glutamine and allowed to recover for 2 days in stationary T75 cell culture flasks. Cells were counted on a Vi-CELL XR (Beckman Coulter), seeded for selection in CD OptiCHO medium lacking glutamine at 1×10^6–2×10^6 cells ml^{-1} and placed in suspension cell culture. Approximately every 2–4 days, pools were monitored for viability and cell density and selective medium were refreshed. Cells were maintained at 36.7 °C in 5% CO_2 humidified incubators. Shake flasks (Corning, VWR and Thomson) were stored on platforms at a shaking speed of 150 r.p.m. with 50-m orbital diameter while 50-l vented cap tubes (Corning) and 24-deep-well plates (VWR) were stored on platforms at a shaking speed of 220 r.p.m. with 50-mm orbital diameter.

Cells were considered recovered from selection once they demonstrated >90% viability and >1×10^6 cells ml^{-1} viable cell density, and were seeded into a 24-deep-well plate fed-batch production assay. Each variant pool was seeded in quadruplicate at 1×10^6 cells ml^{-1} in 3 ml by centrifugation and resuspended in 80% proprietary BAK004-026 and 20% CD OptiCHO media per well. Cells were counted on days 3, 6, 8 and 10 with Guava easyCyte (MilliporeSigma). Cell counts were analyzed using Guava CytoSoft Data Acquisition and Analysis Software. On days 3, 6 and 8, cells were fed 5% of starting volume with CellBoost 7 A (Hyclone) and 0.5% of starting volume with CellBoost 7B (Hyclone). Glucose levels were measured on feed days and fed to maintain a minimum of 8 g l^{-1}. On day 10 the supernatants were clarified with 24-well, 0.2-µm filter plates (Thomson) and replicates were pooled, purified by ProteinA resin and buffer exchanged into PBS at 1 mg ml^{-1} for further analysis.

Statistical analyses

t-tests, Wilcoxon rank-sum tests, Wilcoxon matched-pairs tests, Kolmogorov–Smirnov tests, analysis of variance, correlation, regression, log-rank survival analysis, univariate and multivariable logistic and linear regressions and Fisher exact tests were performed using GraphPad Prism. Bootstrap analyses were performed using R (cran.r-project.org). Benjamini–Hochberg false discovery rate analysis was performed using the function p.adjust (p.values.list, method = 'BH') from R. The number of observations in each analysis is indicated by *n* in both figures and text. Significance is defined as $P < 0.05$. Error bars on figures are standard errors unless otherwise indicated.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Accession numbers for all paired heavy- and light-chain IgG sequences recombinantly expressed were provided through BankIt: 2749610: OR662637–OR663656. The entire set of unique natively paired IgG sequences ($n = 28{,}672$) from PBs ($n = 32{,}948$) of RTS,S vaccines ($n = 45$) are available at https://zenodo.org/records/10019777. Requests for other datasets generated and/or analyzed in the current study will be promptly reviewed by the corresponding authors (malaria@biosimplify.com or kwilliams@atreca.com), and a Material Transfer Agreement (MTA) provided should the request be subject to intellectual property obligations. Materials subject to an MTA will be released pending execution. All other data/materials not subject to an MTA will be provided within a reasonable time frame following the initial request. Data from the clinical trial of RTS,S are reported elsewhere[24]. Source data are provided with this paper.

Code availability

The custom code used in R for bootstrap analyses is available at https://github.com/expositum/RTS_S_CSPBinding_vs_Protection. Custom code used in the TitrationAnalysis tool for SPR analyses is available at https://github.com/DukeCHSI/TitrationAnalysis. All materials used in the study are commercially available except for PBMC samples obtained from the clinical trial (NCT01857869). Antibodies can be expressed as recombinant proteins via gene synthesis and expression using IgG sequence data and methods (see paired IgG sequence repertoire datasets (above) and 'Recombinant antibody production').

References

78. Tan, Y.-C. et al. Sequencing antibody repertoires provides evidence for original anti-genic sin shaping the antibody response to influenza vaccination. *Clin. Immunol.* **151**, 55–65 (2014).
79. Volpe, J. M., Cowell, L. G. & Kepler, T. B. SoDA: implementation of a 3D alignment algorithm for inference of antigen receptor recombinations. *Bioinformatics* **22**, 438–444 (2006).
80. Giudicelli, V. IMGT/GENE-DB: a comprehensive database for human and mouse immunoglobulin and T cell receptor genes. *Nucleic Acids Res.* **33**, D256–D261 (2004).
81. Schwenk, R. et al. IgG2 antibodies against a clinical grade *Plasmodium falciparum* CSP vaccine antigen associate with protection against transgenic sporozoite challenge in mice. *PLoS ONE* **9**, e111020 (2014).
82. Li, K., Horn, G. Q., Alam, S. M., Tomaras, G. D. & Dennison, S. M. Titration analysis: a tool for high-throughput analysis of binding kinetics data for multiple label-free platforms. *Biophys. J.* **120**, 265a–266a (2021).
83. Bruña-Romero, O. et al. Detection of malaria liver-stages in mice infected through the bite of a single *Anopheles* mosquito using a highly sensitive real-time PCR. *Int. J. Parasitol.* **31**, 1499–1502 (2001).
84. Liu, Y. et al. High-throughput screening for developability during early-stage antibody discovery using self-interaction nanoparticle spectroscopy. *mAbs* **6**, 483–492 (2014).
85. Kohli, N. et al. A novel screening method to assess developability of antibody-like molecules. *mAbs* **7**, 752–758 (2015).
86. Honegger, A. & Plückthun, A. Yet another numbering scheme for immunoglobulin variable domains: an automatic modeling and analysis tool. *J. Mol. Biol.* **309**, 657–670 (2001).
87. Rogers, R. S. et al. A view on the importance of 'Multi-Attribute Method' for measuring purity of biopharmaceuticals and improving overall control strategy. *AAPS J.* **20**, 7 (2017).
88. Ong, E.-C., Smidt, P. & McGrew, J. T. Limiting the metabolic burden of recombinant protein expression during selection yields pools with higher expression levels. *Biotechnol. Prog.* **35**, e2839 (2019).
89. Bodwell, J., Swiff, F. & Richardson, J. Long duration electroporation for achieving high level expression of glucocorticoid receptors in mammalian cell lines. *J. Steroid Biochem. Mol. Biol.* **68**, 77–82 (1999).

Acknowledgements

We thank J. Kirchner for scientific, business and leadership contributions and for manuscript review and editing; L. M. Stuart for scientific, business and leadership contributions; L. Shackelton for scientific, business and project management contributions; D. Hopkins for project management contributions; K. Andrews and H. Liu for scientific and project management contributions; R. Scott Miller for manuscript review and editing; T. A. Serafini for corporate support and business and scientific contributions; S. E. Gould for manuscript review, editing, figure improvements and scientific contributions;

S. Mudrak for project management support, manuscript review and editing; R. H. C. Huntwork and G. Q. Horn for binding analyses of engineered variants; and A. Birkett and C. Gerard for manuscript review and editing. F.Z., Y.F.-G. and S.M.-T. thank the Insectary Core of the Malaria Research Institute and the Bloomberg Philanthropies for continued support. This work was supported by the Bill and Melinda Gates Foundation (BMGF) at Atreca, Inc. and at Just—Evotec Biologics, Inc. (no. INV-008062), with grants to F.Z. (INV-001763) and G.D.T. (nos. OPP1151372 and INV008612), and by PATH-MVI with a grant to F.Z. BMGF contributed to experimental design and interpretation of results. We also acknowledge editorial assistance from M. Hyde and J. McGuire at BOLDSCIENCE, Inc. The contents of this publication are solely the responsibility of the authors and do not represent the official views of BMGF.

Author contributions

Conceptualization was undertaken by K.L.W. and D.E.E. Methodology was the responsibility of K.L.W., S.G., Y.F.-G., D.K., K.S.W., C.S., P.S., S.Z.J., K.L., S.M.D., S.M.-T., X.C., U.W.-R., M.W., E.J., C.R.K., C.O., J.G., J.E.M., J.A.R., Y.C.T., G.C., S.M.L., W.H.R., G.D.T., F.Z., R.R.K. and D.E.E. Software was the responsibility of S.G. K.L.W., S.G., Y.F.-G., D.K., K.S.W., C.S., P.S., S.Z.J., K.L., R.S.M., S.D., R.R.K. and D.E.E. performed validation. Formal analysis was undertaken by K.L.W., S.G., Y.F.-G., D.K., K.S.W., C.S., P.S., S.Z.J., K.L., S.M.D., U.W.-R., M.W., R.S.M., C.O., J.G., S.D., Y.C.T., S.M.L., W.H.R., G.D.T., F.Z., R.R.K. and D.E.E. K.L.W., S.G., Y.F.-G., C.S., P.S., S.D., G.D.T., F.Z., R.R.K. and D.E.E. were responsible for resources. Data curation was performed by K.L.W., S.G., Y.F.-G., D.K., K.S.W., C.S., P.S., S.Z.J., K.L., R.S.M., S.D., R.R.K. and D.E.E. K.L.W. and D.E.E. wrote the original draft. Review and editing of the paper were undertaken by K.L.W., D.K., K.S.W., C.S., P.S., S.Z.J., K.L., S.M.D., M.W., E.J., R.S.M., C.R.K., C.O., J.G., G.C., S.M.L., S.D., G.D.T., F.Z., R.R.K. and D.E.E. K.L.W. and D.E.E. administered the project.

Competing interests

W.H.R. owns equity in, serves as a consultant to and is a member of the Board of Directors of Atreca, Inc. K.L.W., S.G., K.S.W. and S.M.L. own equity in, and are employed by, Atreca, Inc. D.E.E. and M.W. own equity in, and serve as consultants to, Atreca, Inc. G.C. and D.K. own equity in Atreca, Inc. C.S., P.S., S.Z.J. and R.R.K. are employed by Just–Evotec Biologics. E.J. owns equity in, and is employed by, GSK. U.W.-R. is employed by BioNTech. J.G. owns equity in, is employed by and is a member of the Board of Directors of Centivax, Inc. Y.C.T. owns equity in Atreca, Inc. and is employed by Nuevocor Pte. Ltd. K.L.W., S.M.L., R.R.K. and D.E.E. are coinventors on patent applications filed by Atreca, Inc. that include antibodies to CSP. The remaining authors declare no competing interests.

Additional information

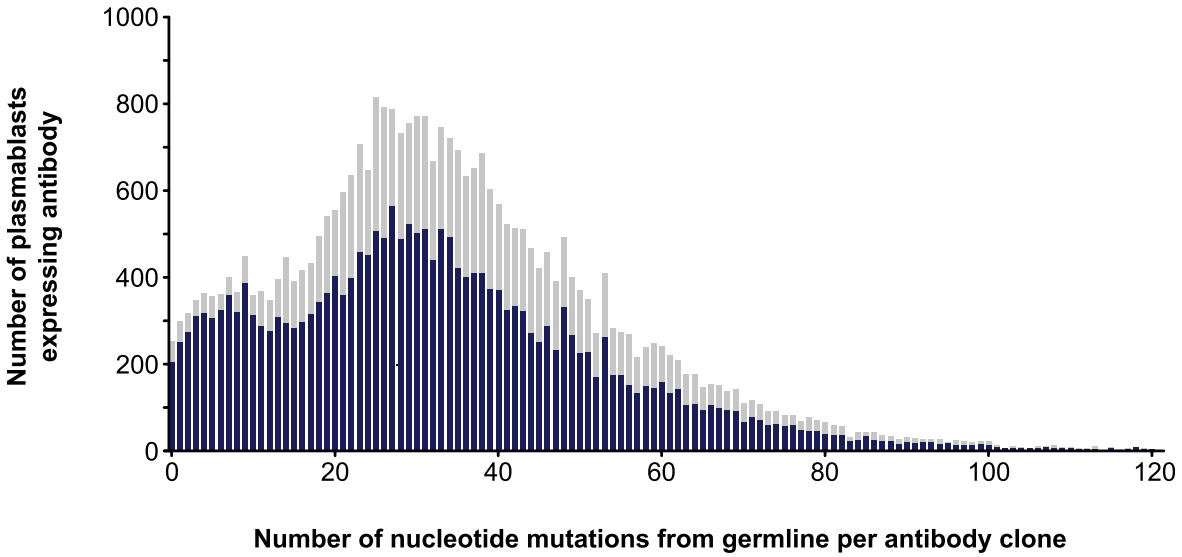
Extended data is available for this paper at https://doi.org/10.1038/s41591-023-02659-z.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41591-023-02659-z.

Correspondence and requests for materials should be addressed to Katherine L. Williams or Daniel E. Emerling.

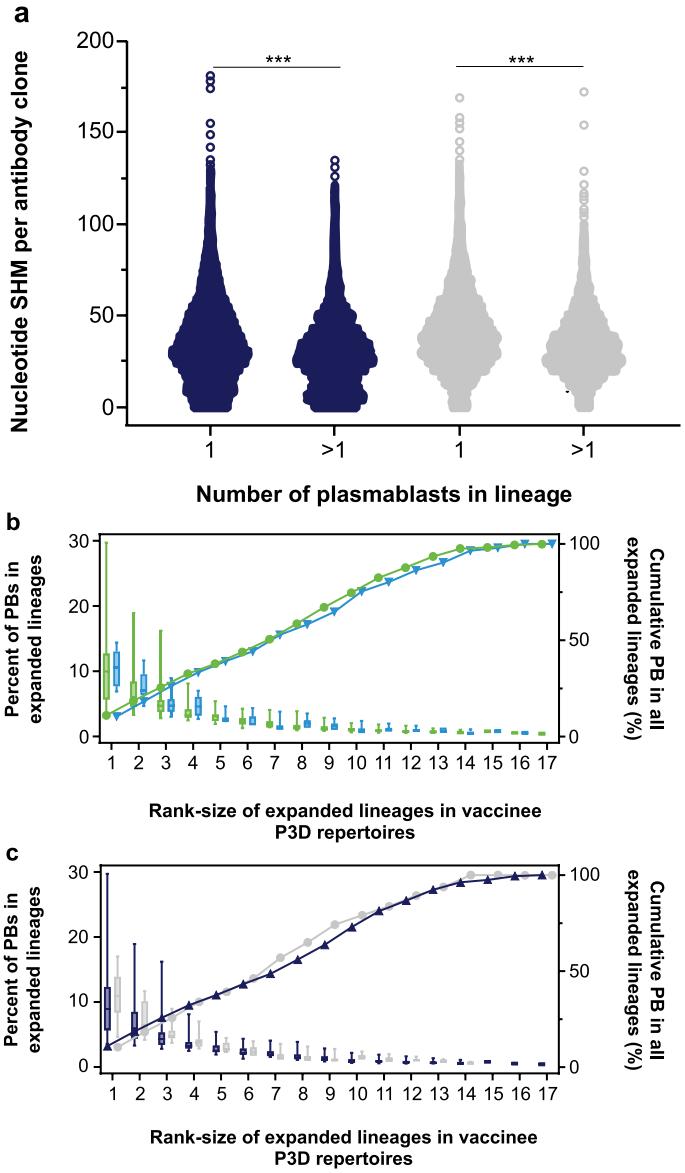
Peer review information *Nature Medicine* thanks Joseph Jardine and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Alison Farrell, in collaboration with the *Nature Medicine* team.

https://doi.org/10.1038/s41591-023-02659-z



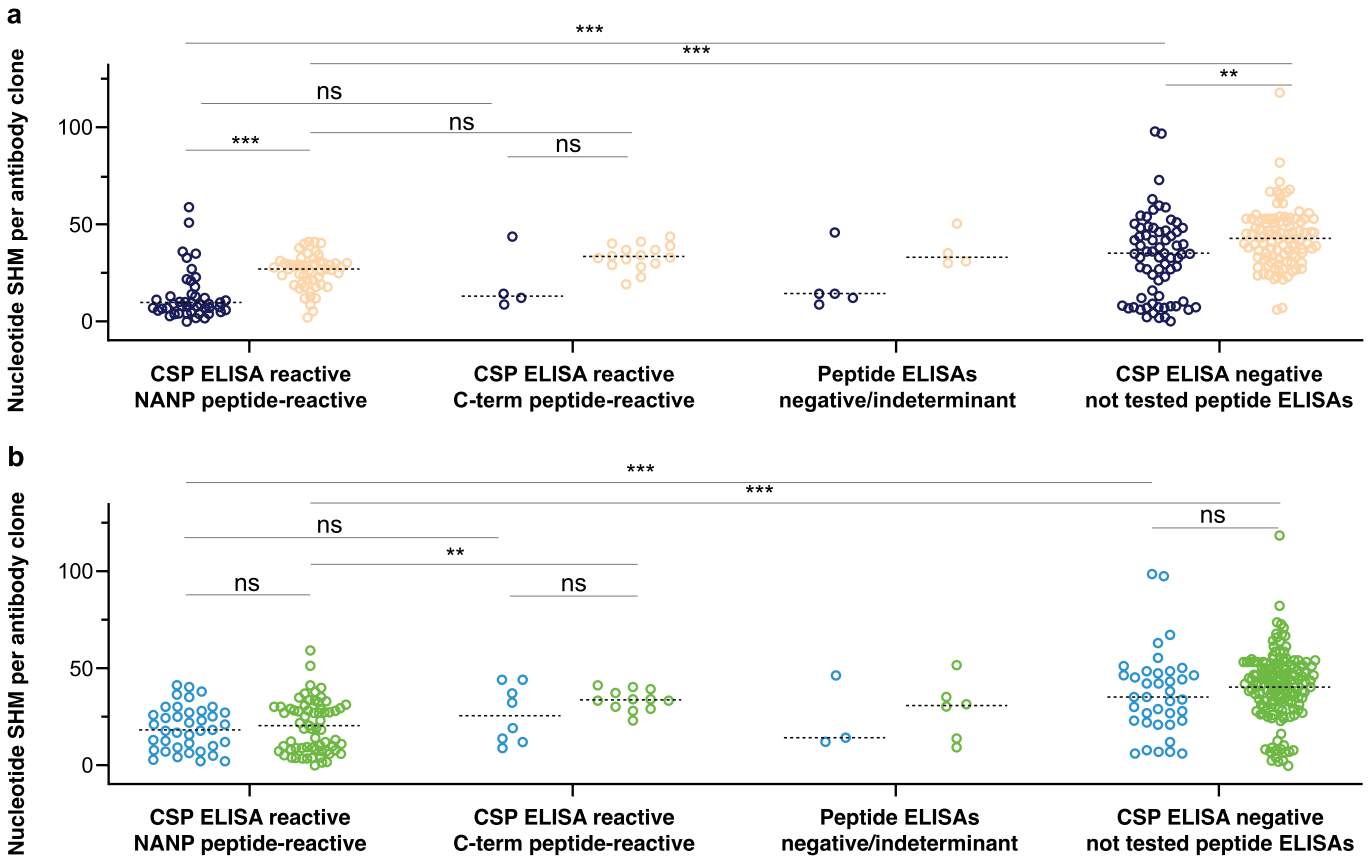
Extended Data Fig. 1 | Levels of SHM among antibody sequences. Histogram of the number of nucleotide mutations from germline (SHM) of combined IgG heavy and light chains from blood PBs collected 7 days after administration of the third dose (blue, $n = 22,319$) or fourth dose (grey, $n = 10,629$) of RTS,S.

https://doi.org/10.1038/s41591-023-02659-z



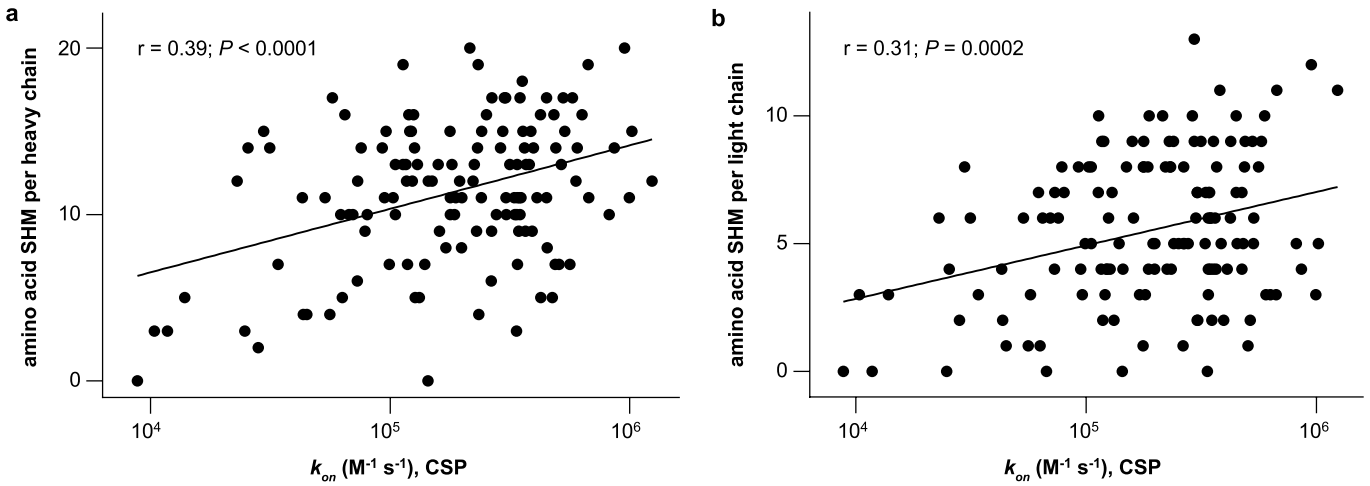
Extended Data Fig. 2 | SHM and size distribution of expanded PB IgG lineages. **a**, Range of nucleotide mutations from germline (SHM) in heavy plus light chain for antibodies in expanded lineages (>1 PB) compared to antibodies in lineages that lack evidence of recent expansion (1 PB). PBs were collected 7 days after administration of the third dose (blue, $n = 11,478$ and $n = 10,841$, respectively) or fourth dose (grey, $n = 5,538$ and $n = 5,091$, respectively) of RTS,S, ***$P < 0.0001$ unpaired two-tailed Mann–Whitney test. **b**, **c**, By vaccinee, the size of each expanded lineage was calculated by dividing the number of PBs in that lineage by the number of PBs in all expanded lineages within each repertoire and then assigning a rank-size. Box and whisker plots represent the distribution of lineage sizes across vaccinees for each rank-size from, **b**, protected (green, $n = 36$) versus unprotected (blue, $n = 9$) vaccinees, and, **c**, from vaccinees administered a third standard dose (012 M, grey, $n = 15$) versus vaccinees administered a delayed fractional dose (Fx017M, blue, $n = 30$) of RTS,S. Boxes indicate interquartile ranges, lines within boxes are medians, and whiskers represent minimum and maximum across vaccinees.

https://doi.org/10.1038/s41591-023-02659-z



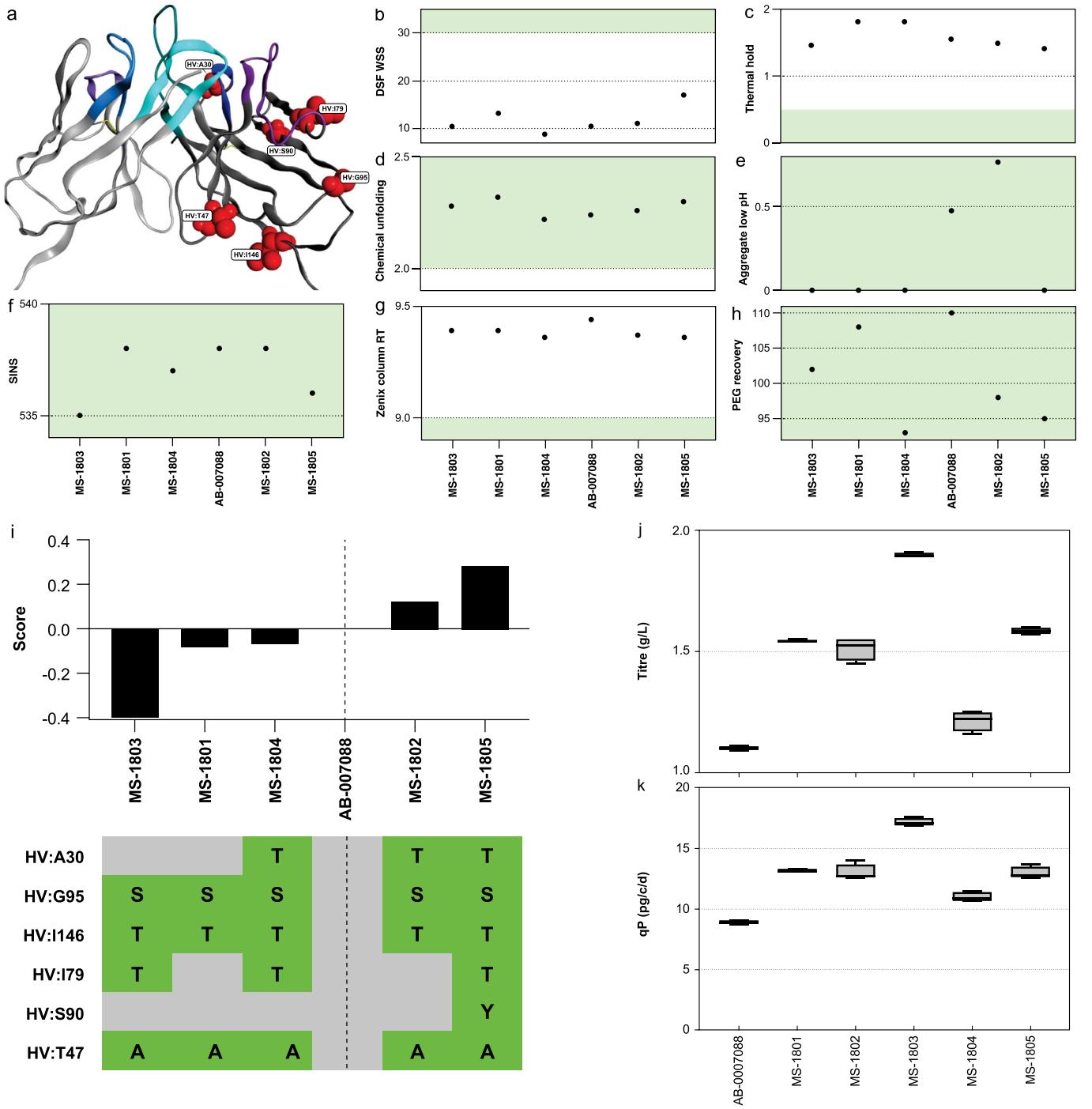
Extended Data Fig. 3 | SHM and CSP-peptide binding of mAbs versus RTS,S dose group and vaccinee protection status. a, **b**, Number of nucleotide mutations from germline (SHM) for mAbs from expanded lineages of vaccinees, **a**, administered a third standard dose (012 M, dark blue, $n = 15$) or a delayed fractional dose (Fx017M, orange, $n = 30$) of RTS,S with CSP-reactive mAbs shown that are reactive to the repeat region peptide (NANP6, $n = 45$ and $n = 53$, respectively), to the C-terminal region peptide (C-terminal, $n = 4$ and $n = 16$, respectively), or to neither peptide (negative, $n = 5$ and $n = 4$, respectively) in comparison to mAbs not reactive in the CSP ELISA ($n = 72$ and $n = 113$, respectively), and, **b**, from vaccinees that were protected (green, $n = 36$) or unprotected (light blue, $n = 9$) with mAbs shown that are reactive to repeat region peptide ($n = 59$ and $n = 39$, respectively)], to the C-terminal region peptide ($n = 12$ and $n = 8$, respectively), or to neither peptide (negative, $n = 6$ and $n = 3$, respectively) in comparison to mAbs not reactive in the CSP ELISA ($n = 147$ and $n = 38$, respectively), lines are medians, $^{***}P < 0.0001$, $^{**}P < 0.001$, $^{*}P < 0.02$, or $P > 0.1$ ('ns'), unpaired two-tailed Mann–Whitney test.

https://doi.org/10.1038/s41591-023-02659-z



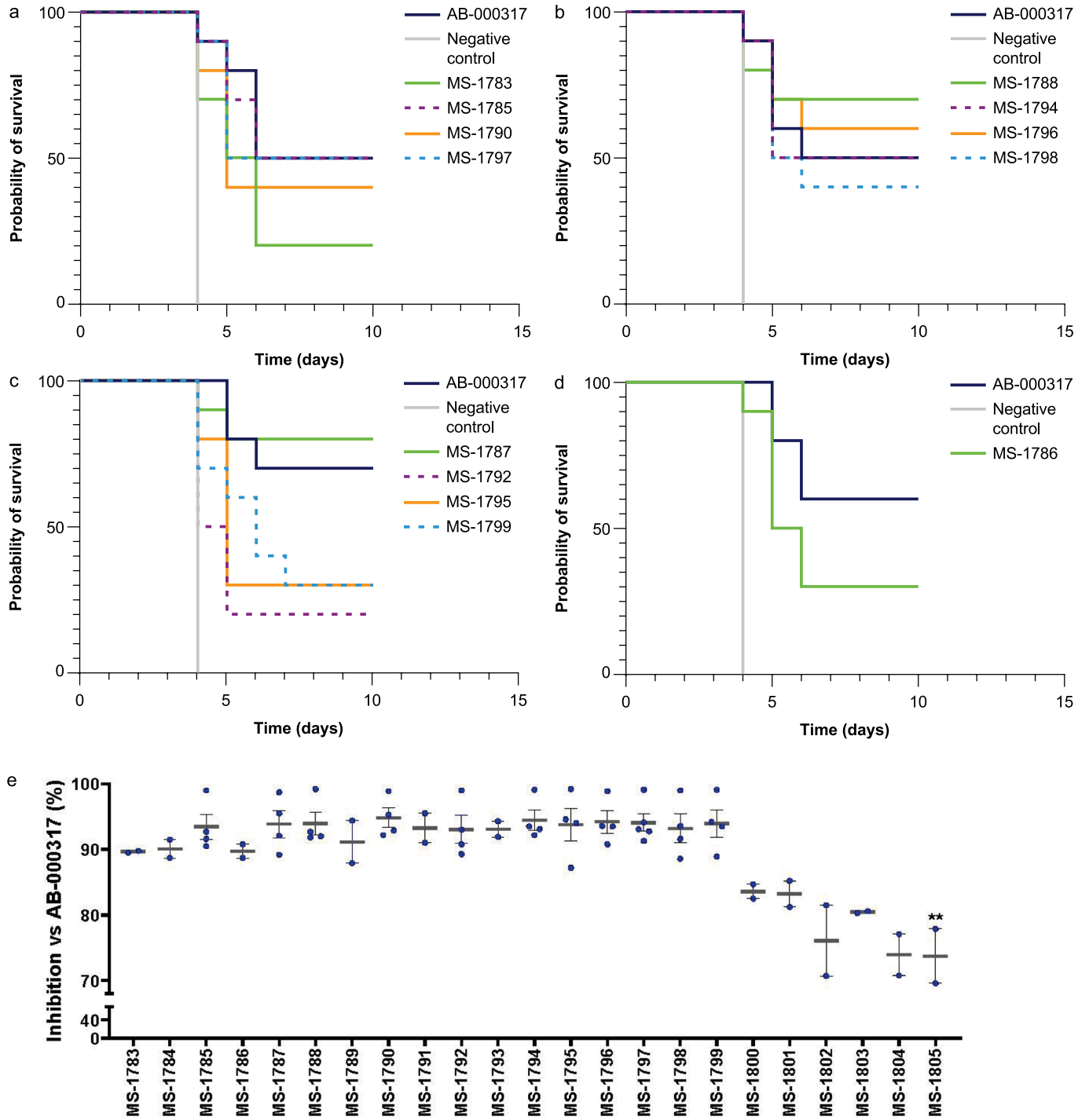
Extended Data Fig. 4 | CSP-binding on-rates and amino acid divergence from germline. a, **b**, Simple two-tailed linear regression of log-transformed SPR binding on-rates (k_{on}) to CSP of mAbs ($n = 141$) from 35 of the most efficacious lineages compared to the number of, **a**, heavy, and, **b**, light chain amino acid mutations from germline (SHM). For correlations of non-transformed data, see Extended Data Table 1.

https://doi.org/10.1038/s41591-023-02659-z



Extended Data Fig. 5 | Generation and developability properties of engineered variants from backup molecule. a, Mutations made to generate engineered variants are depicted in the Fv region of a Fab structure model of AB-007088 built in the MOE Antibody Modeler (*see* Methods). The light chain framework is in silver and the heavy chain in dim grey. The CDRs for light and heavy chains are indicated, respectively, as CDR1 (light and dark blue), CDR2 (light and dark purple), and CR3 (light and dark cyan). Mutation sites of engineered variants are labelled as stability violations (red, *see* Methods). **b**–**h**, Developability properties of engineered variants of parental backup, AB-007088, as characterized by scores from, **b**, DSF WSS, **c**, thermal hold, **d**, chemically induced unfolding, **e**, low pH stability (aggregate low pH hold), **f**, self-interaction nanoparticle spectroscopy (SINS), **g**, stand-up monolayer affinity chromatography (Zenix column RT), **h**, relative solubility analysis (PEG recovery) with green shading that indicates preferred ranges. **i**, Aggregate score of assay panel results from **b**-**h** for each engineered variant with mutations shown in comparison to AB-007088. **j**–**k**, Characterization of stable cell production pools generated from engineered variants compared to AB-007088 for, **j**, production titres, (Titre, g/L) and, **k**, cell-specific productivity (qP, pg per cell per day). Boxes indicate interquartile ranges, lines within boxes are medians, and whiskers represent minimum and maximum (4 independent replicates of each antibody).

https://doi.org/10.1038/s41591-023-02659-z



Extended Data Fig. 6 | In vivo activity of engineered mAb variants. a–d, Mosquito-bite parasitaemia survival curves for 13 variants ($n = 10$ mice), including the eight that had the best aggregate score from the developability assays. **e**, Sporozoite liver burden percent inhibition compared to untreated, infected mice (geometric mean, $n = 5$) and normalized to the activity of AB-000317, lines and bars indicate mean ± SEM, 2–5 independent experiments per mAb, $^{**}P = 0.005$ or $P > 0.05$ for all other variants, two-tailed Dunn's multiple comparisons test verses AB-000317 (see Extended Data Table 3 for details).

https://doi.org/10.1038/s41591-023-02659-z

Extended Data Table 1 | Correlation test results comparing SHM, CSP-peptide binding data, and liver burden inhibition data

			Heavy chain				Light chain					
			SHM (nucleotide)		SHM (amino acid)		SHM (nucleotide)		SHM (amino acid)		LB Inhibition[c]	
Binding peptide in SPR	Binding[a]	Test[b]	*p*	r	*p*	r	*p*	r	*p*	r	*p*	r
CSP (3D7)	K_{off}	S	>0.8	-0.02	>0.7	0.03	>0.08	-0.17	>0.07	-0.17	>0.2	-0.18
CSP (3D7)	K_{off}	P	>0.3	-0.09	>0.7	-0.03	>0.1	-0.14	>0.1	-0.14	**<0.01**	**-0.37**
CSP (3D7)	K_{off}	LR	>0.3	-0.09	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	**<0.03**	**-0.32**
NPNA3	K_{off}	S	**<0.0001**	**-0.48**	**<0.0001**	**-0.52**	**<0.0001**	**-0.34**	**<0.0002**	**-0.32**	**<0.0001**	**-0.48**
NPNA3	K_{off}	P	**<0.0001**	**-0.34**	**<0.0004**	**-0.30**	**<0.03**	**-0.19**	**<0.05**	**-0.17**	**<0.0001**	**-0.53**
NPNA3	K_{off}	LR	**<0.0001**	**-0.48**	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	**<0.0001**	**-0.60**
Junction	K_{off}	S	**<0.03**	**-0.26**	**<0.04**	**-0.25**	**<0.003**	**-0.36**	**<0.006**	**-0.33**	**<0.05**	**-0.31**
Junction	K_{off}	P	**<0.03**	**-0.27**	**<0.04**	**-0.25**	**<0.004**	**-0.35**	**<0.004**	**-0.35**	**<0.005**	**-0.43**
Junction	K_{off}	LR	**<0.05**	**-0.24**	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	**<0.03**	**-0.35**
NPDPNANP2NVDP	K_{off}	S	**<0.002**	**-0.30**	**<0.002**	**-0.30**	>0.06	-0.18	>0.09	-0.16	>0.07	-0.24
NPDPNANP2NVDP	K_{off}	P	**<0.04**	**-0.20**	**>0.07**	**-0.17**	>0.1	-0.16	>0.2	-0.12	**<0.003**	**-0.38**
NPDPNANP2NVDP	K_{off}	LR	**<0.002**	**-0.30**	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	**<0.005**	**-0.37**
NVDP3NANP2	K_{off}	S	**<0.02**	**-0.24**	**<0.05**	**-0.18**	**<0.04**	**-0.20**	**<0.05**	**-0.18**	**<0.002**	**-0.40**
NVDP3NANP2	K_{off}	P	>0.4	0.06	>0.2	0.10	>0.6	-0.04	>0.9	-0.01	>0.3	-0.13
NVDP3NANP2	K_{off}	LR	>0.05	-0.18	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	**<0.0007**	**-0.42**
NANP6	K_{off}	S	>0.3	-0.11	>0.6	-0.05	>0.7	0.04	>0.2	0.14	>0.3	0.15
NANP6	K_{off}	P	>0.06	-0.20	>0.1	-0.15	>0.1	-0.16	>0.1	-0.15	>0.6	0.09
NANP6	K_{off}	LR	>0.2	-0.12	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	>0.7	0.05
CSP (3D7)	K_{on}	S	**<0.008**	**0.22**	**<0.0008**	**0.28**	**<0.02**	**0.20**	**<0.008**	**0.22**	>0.4	0.09
CSP (3D7)	K_{on}	P	**<0.004**	**0.25**	**<0.0007**	**0.28**	**<0.02**	**0.21**	**<0.007**	**0.23**	>0.4	0.09
CSP (3D7)	K_{on}	LR	**<0.0005**	**0.42**	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	>0.5	0.07
NPNA3	K_{on}	S	**<0.0001**	**0.47**	**<0.0001**	**0.53**	**<0.007**	**0.23**	**<0.02**	**0.20**	**<0.04**	**0.25**
NPNA3	K_{on}	P	**<0.0001**	**0.40**	**<0.0001**	**0.47**	**<0.05**	**0.17**	>0.07	0.15	**<0.02**	**0.31**
NPNA3	K_{on}	LR	**<0.0001**	**0.59**	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	**<0.04**	**0.26**
Junction	K_{on}	S	**<0.004**	**0.35**	**<0.0008**	**0.40**	**<0.008**	**0.32**	**<0.005**	**0.34**	>0.05	0.30
Junction	K_{on}	P	**<0.03**	**0.27**	**<0.008**	**0.32**	**<0.03**	**0.28**	**<0.004**	**0.35**	>0.09	0.26
Junction	K_{on}	LR	**<0.003**	**0.45**	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	**<0.04**	**0.32**
NPDPNANP2NVDP	K_{on}	S	**<0.0001**	**0.40**	**<0.0001**	**0.44**	>0.1	0.13	>0.06	0.17	>0.1	0.20
NPDPNANP2NVDP	K_{on}	P	**<0.001**	**0.31**	**<0.0002**	**0.36**	>0.3	0.10	>0.1	0.13	**<0.05**	**0.27**
NPDPNANP2NVDP	K_{on}	LR	**<0.0009**	**0.43**	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	>0.06	0.24
NVDP3NANP2	K_{on}	S	**<0.003**	**0.27**	**<0.0002**	**0.33**	**<0.02**	**0.22**	**<0.02**	**0.22**	>0.4	0.09
NVDP3NANP2	K_{on}	P	**<0.004**	**0.26**	**<0.0004**	**0.31**	>0.1	0.13	>0.1	0.14	>0.8	0.03
NVDP3NANP2	K_{on}	LR	**<0.002**	**0.38**	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	>0.5	0.08
NANP6	K_{on}	S	**<0.0009**	**0.28**	**<0.0001**	**0.34**	**<0.04**	**0.18**	>0.05	0.16	>0.8	-0.02
NANP6	K_{on}	P	**<0.009**	**0.22**	**<0.0004**	**0.29**	**<0.03**	**0.19**	**<0.04**	**0.18**	>0.7	0.05
NANP6	K_{on}	LR	**<0.005**	**0.34**	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	>0.7	-0.04
CSP (3D7)	K_D	S	**<0.004**	**-0.27**	**<0.006**	**-0.26**	**<0.0003**	**-0.34**	**<0.0005**	**-0.33**	>0.2	-0.17
CSP (3D7)	K_D	P	>0.06	-0.18	>0.1	-0.14	**<0.03**	**-0.22**	**<0.03**	**-0.22**	>0.1	-0.20
CSP (3D7)	K_D	LR	**<0.002**	**-0.46**	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	>0.1	-0.22
NPNA3	K_D	S	**<0.0001**	**-0.53**	**<0.0001**	**-0.57**	**<0.0001**	**-0.33**	**<0.0004**	**-0.30**	**<0.0003**	**-0.43**
NPNA3	K_D	P	**<0.002**	**-0.27**	**<0.004**	**-0.25**	**<0.03**	**-0.20**	**<0.03**	**-0.20**	**<0.003**	**-0.36**
NPNA3	K_D	LR	**<0.0001**	**-0.70**	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	**<0.0001**	**-0.54**
Junction	K_D	S	**<0.008**	**-0.32**	**<0.005**	**-0.34**	**<0.005**	**-0.34**	**<0.005**	**-0.34**	>0.06	-0.29
Junction	K_D	P	**<0.004**	**-0.35**	**<0.002**	**-0.39**	**<0.005**	**-0.34**	**<0.005**	**-0.34**	**<0.05**	**-0.31**
Junction	K_D	LR	**<0.009**	**-0.40**	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	**<0.02**	**-0.38**
NPDPNANP2NVDP	K_D	S	**<0.0001**	**-0.38**	**<0.0001**	**-0.40**	>0.06	-0.18	**<0.05**	**-0.19**	>0.08	-0.22
NPDPNANP2NVDP	K_D	P	**<0.02**	**-0.23**	**<0.02**	**-0.23**	**<0.05**	**-0.19**	>0.1	-0.16	**<0.02**	**-0.33**
NPDPNANP2NVDP	K_D	LR	**<0.002**	**-0.40**	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	**<0.009**	**-0.34**
NVDP3NANP2	K_D	S	**<0.0002**	**-0.34**	**<0.0002**	**-0.34**	**<0.002**	**-0.28**	**<0.003**	**-0.27**	**<0.002**	**-0.40**
NVDP3NANP2	K_D	P	>0.3	0.09	>0.2	0.10	>0.6	-0.04	>0.8	-0.02	>0.3	-0.11
NVDP3NANP2	K_D	LR	**<0.0005**	**-0.43**	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	**<0.006**	**-0.35**
NANP6	K_D	S	**<0.0005**	**-0.37**	**<0.0009**	**-0.35**	**<0.04**	**-0.22**	>0.4	-0.09	>0.2	0.18
NANP6	K_D	P	**<0.002**	**-0.33**	**<0.003**	**-0.32**	**<0.02**	**-0.27**	**<0.008**	**-0.28**	>0.4	0.13
NANP6	K_D	LR	>0.1	-0.27	*nd*	*nd*	*nd*	*nd*	*nd*	*nd*	>0.5	0.10
Functional activity *in vivo* vs. SHM		**Test[b]**	***p***	**r**	***p***	**r**	***p***	**r**	***p***	**r**		
Liver burden inhibition[c]		S	**<0.02**	**0.30**	**<0.0006**	**0.41**	**<0.0002**	**0.44**	**<0.0001**	**0.50**	NA	NA
Liver burden inhibition[c]		P	**<0.0005**	**0.41**	**<0.0001**	**0.48**	**<0.0004**	**0.42**	**<0.0001**	**0.47**	NA	NA
Liver burden inhibition[c]		LR	**<0.0006**	**0.41**	**<0.0001**	**0.48**	**<0.0004**	**0.42**	**<0.0001**	**0.47**	NA	NA

nd, not done; NA, not applicable; LB, liver burden mouse model; NPNA3, NPNANPNANPNA; Junction, KQPADGNPDPNANPN NPDPNANP2NVDP, NPDPNANPNVDPNANP; NVDP3NANP2, NVDPNANPNVDPNANPNVDP; NANP6, NANPNANPNANPNANPNANPNANP

[a] SPR determined k_{on} (k_a), K_{off} (k_d), and K_D, see methods

[b] S, Spearman correlation; P, Pearson correlation; LR, simple linear regression using log transformedK_{off}, K_{on}, or KD data, two-tailed

[c] liver burden inhibition normalized to percent inhibition of AB-000317 tested in parallel

https://doi.org/10.1038/s41591-023-02659-z

Extended Data Table 2 | Binding data to CSP-derived peptides of engineered variant mAbs

	NPNA3				NVDP3NANP2			
mAB ID	K_D, avg. (M)	SEM	Fold-difference vs. parental	compared to partental mAb[a]	K_D, avg. (M)	SEM	Fold-difference vs. parental	compared to partental mAb[a]
AB-000224	3.0E-09	4.6E-10	NA	NA	1.3E-09	5.4E-11	NA	NA
MS-1783	6.6E-09	4.5E-10	0.45	$P<0.05$	1.9E-09	3.3E-10	0.66	$P>0.05$
MS-1784	5.4E-09	5.1E-11	0.56	$P<0.05$	9.8E-10	1.1E-10	1.3	$P>0.05$
MS-1785	5.7E-09	2.4E-10	0.52	$P<0.05$	9.5E-10	2.1E-11	1.4	$P>0.05$
MS-1786	5.7E-09	2.6E-10	0.53	$P<0.05$	9.1E-10	1.2E-10	1.4	$P>0.05$
MS-1787	5.7E-09	3.4E-10	0.52	$P<0.05$	1.1E-09	3.5E-11	1.1	$P>0.05$
MS-1788	4.4E-09	2.1E-10	0.68	$P>0.05$	1.2E-09	1.6E-10	1.1	$P>0.05$
MS-1789	4.5E-09	1.9E-10	0.66	$P>0.05$	8.7E-10	4.4E-11	1.5	$P>0.05$
MS-1790	4.9E-09	9.4E-10	0.61	$P>0.05$	1.0E-09	2.0E-10	1.3	$P>0.05$
MS-1791	4.6E-09	8.0E-10	0.65	$P>0.05$	1.0E-09	1.5E-10	1.2	$P>0.05$
MS-1792	4.5E-09	1.2E-09	0.66	$P>0.05$	9.5E-10	1.8E-10	1.4	$P>0.05$
MS-1793	4.7E-09	1.1E-09	0.63	$P>0.05$	1.1E-09	1.1E-10	1.1	$P>0.05$
MS-1794	2.8E-09	1.6E-10	1.1	$P>0.05$	1.2E-09	6.7E-11	1.1	$P>0.05$
MS-1795	4.8E-09	1.2E-10	0.62	$P>0.05$	1.1E-09	9.1E-11	1.1	$P>0.05$
MS-1796	4.4E-09	1.7E-10	0.68	$P>0.05$	1.2E-09	3.8E-10	1.1	$P>0.05$
MS-1797	5.0E-09	2.0E-10	0.60	$P>0.05$	6.8E-10	1.0E-10	1.9	$P>0.05$
MS-1798	3.8E-09	4.4E-11	0.78	$P>0.05$	7.2E-10	2.8E-11	1.8	$P>0.05$
MS-1799	2.4E-09	7.7E-11	1.3	$P>0.05$	9.1E-10	8.3E-11	1.4	$P>0.05$
AB-007088	1.9E-09	1.0E-10	NA	NA	9.4E-10	1.8E-10	NA	NA
MS-1801	3.1E-09	2.2E-10	0.61	$P>0.05$	1.3E-09	2.5E-10	0.73	$P>0.05$
MS-1802	8.3E-09	2.6E-10	0.23	$P<0.05$	8.3E-10	9.9E-11	1.1	$P>0.05$
MS-1803	2.6E-09	4.6E-10	0.74	$P>0.05$	1.2E-09	2.5E-10	0.80	$P>0.05$
MS-1804	6.3E-09	7.8E-10	0.30	$P<0.05$	5.7E-10	1.5E-10	1.6	$P>0.05$
MS-1805	4.0E-09	4.5E-10	0.48	$P<0.05$	8.0E-10	1.0E-10	1.2	$P>0.05$

NPNA3: NPNANPNANPNA; NVDP3NANP2: NVDPNANPNVDPNANPNVDP

NA, not applicable

[a] ANOVA and Dunnett's multiple comparisons test, two-tailed

https://doi.org/10.1038/s41591-023-02659-z

Extended Data Table 3 | Developability ranking and in vivo pharmacology of engineered variant mAbs

Model	Experiment	Parental mAb	mAb ID	Rank, developability[a]	LB Inhibition (%)[b]	Dose (µg)	Versus to AB-000317[c]	Versus parental mAb[c]	HR[d]	95% CI[d]	$[mAb]_{sera}$ (µg/ml)	Std dev $[mAb]_{sera}$	Versus $[AB\text{-}000317]_{sera}$ (%)	Versus AB-000317[c]	Versus parental mAb[c]
LB	225	NA	AB-000317	NA	NA	100	NA	NA	NA	NA	64.7	2.6	100.0	NA	NA
LB	225	NA	AB-000224	16	103.6	100	>0.5	NA	NA	NA	68.7	5.4	106.1	>0.5	NA
LB	225	AB-000224	MS-1786	8	100.5	100	>0.5	>0.5	NA	NA	48.8	6.1	75.5	>0.5	>0.3
LB	225	AB-000224	MS-1788	7	102.7	100	>0.5	>0.5	NA	NA	46.9	4.7	72.5	>0.4	>0.1
LB	225	AB-000224	MS-1789	14	97.3	100	>0.5	>0.1	NA	NA	40.2	2.8	62.1	0.0047	0.0012
LB	225	AB-000224	MS-1791	13	100.7	100	>0.5	>0.5	NA	NA	44.8	4.0	69.3	>0.1	0.0528
LB	225	AB-000224	MS-1793	9	101.8	100	>0.5	>0.5	NA	NA	50.7	7.6	78.3	>0.5	>0.5
LB	225	AB-000224	MS-1794	2	102.2	100	>0.5	>0.5	NA	NA	48.4	2.0	74.7	>0.5	>0.4
LB	225	AB-000224	MS-1799	15	98.5	100	>0.5	>0.2	NA	NA	41.2	3.6	63.6	0.0119	0.0034
LB	228	NA	AB-000317	NA	NA	100	NA	NA	NA	NA	108.0	26.3	100.0	NA	NA
LB	228	AB-000224	MS-1783	3	97.3	100	>0.5	*nd*	NA	NA	72.7	18.8	67.4	>0.5	*nd*
LB	228	AB-000224	MS-1784	11	96.1	100	>0.5	*nd*	NA	NA	77.0	15.6	71.3	>0.5	*nd*
LB	228	AB-000224	MS-1785	4	98.2	100	>0.5	*nd*	NA	NA	66.2	16.9	61.4	>0.1	*nd*
LB	228	AB-000224	MS-1787	6	96.8	100	>0.5	*nd*	NA	NA	46.9	8.1	43.4	0.0006	*nd*
LB	228	AB-000224	MS-1790	4	100.0	100	>0.5	*nd*	NA	NA	66.5	12.9	61.6	>0.2	*nd*
LB	228	AB-000224	MS-1792	8	96.8	100	>0.5	*nd*	NA	NA	47.6	6.3	44.1	0.0011	*nd*
LB	228	AB-000224	MS-1795	12	94.6	100	>0.5	*nd*	NA	NA	55.6	4.0	51.5	0.0207	*nd*
LB	228	AB-000224	MS-1796	10	98.4	100	>0.5	*nd*	NA	NA	91.7	10.0	85.0	>0.5	*nd*
LB	304	NA	AB-000317	NA	NA	100	NA	NA	NA	NA	52.3	4.6	100.0	NA	NA
LB	304	NA	AB-007088	3	89.3	100	>0.4	NA	NA	NA	54.4	11.1	104.1	>0.5	NA
LB	304	AB-000224	MS-1797	1	102.6	100	>0.5	*nd*	NA	NA	32.2	3.7	61.6	0.0175	*nd*
LB	304	AB-000224	MS-1798	5	99.5	100	>0.5	*nd*	NA	NA	30.2	2.9	57.7	0.0014	*nd*
LB	304	AB-007088	MS-1801	5	91.2	100	>0.5	>0.5	NA	NA	37.0	8.1	70.7	>0.2	>0.3
LB	304	AB-007088	MS-1802	2	91.6	100	>0.5	>0.5	NA	NA	40.8	10.4	78.0	>0.5	>0.5
LB	304	AB-007088	MS-1803	6	90.2	100	>0.5	>0.5	NA	NA	34.7	5.1	66.4	0.0943	>0.1
LB	304	AB-007088	MS-1804	4	86.6	100	>0.1	>0.5	NA	NA	38.7	8.2	73.9	>0.5	>0.5
LB	304	AB-007088	MS-1805	1	78.1	100	0.0105	>0.5	NA	NA	34.2	3.9	65.4	0.089	>0.1
LB	325	NA	AB-000317	NA	NA	100	NA	NA	NA	NA	61.5	13.6	100.0	NA	NA
LB	325	NA	AB-000224	16	103.2	100	>0.5	NA	NA	NA	58.1	4.8	94.4	>0.5	NA
LB	325	AB-000224	MS-1783	3	96.3	100	>0.5	0.0603	NA	NA	30.9	2.5	50.2	0.0119	0.0119
LB	325	AB-000224	MS-1784	11	98.4	100	>0.5	>0.2	NA	NA	31.9	2.8	51.8	0.0351	0.0351
LB	325	AB-000224	MS-1785	4	99.7	100	>0.5	>0.5	NA	NA	31.3	2.7	50.9	0.0306	0.0306
LB	325	AB-000224	MS-1786	8	95.4	100	>0.5	0.0404	NA	NA	32.8	4.6	53.3	0.0389	0.0389
LB	325	AB-000224	MS-1787	6	99.1	100	>0.5	>0.4	NA	NA	38.5	9.6	62.6	>0.3	>0.3
LB	325	AB-000224	MS-1788	7	99.0	100	>0.5	>0.5	NA	NA	36.0	5.2	58.6	>0.3	>0.3
LB	325	AB-000224	MS-1789	14	101.6	100	>0.5	>0.5	NA	NA	30.1	2.7	49.0	0.0053	0.0053
LB	328	NA	AB-000317	NA	NA	100	NA	NA	NA	NA	72.7	5.1	100.0	NA	NA
LB	328	AB-000224	MS-1790	4	102.9	100	>0.5	*nd*	NA	NA	41.0	3.5	56.4	0.0067	*nd*
LB	328	AB-000224	MS-1791	13	103.1	100	>0.5	*nd*	NA	NA	42.7	2.8	58.8	0.0306	*nd*
LB	328	AB-000224	MS-1792	8	98.1	100	>0.5	*nd*	NA	NA	37.1	4.5	51.1	0.0002	*nd*
LB	328	AB-000224	MS-1793	9	101.8	100	>0.5	*nd*	NA	NA	45.1	1.9	62.1	>0.4	*nd*
LB	328	AB-000224	MS-1794	2	101.0	100	>0.5	*nd*	NA	NA	41.2	2.2	56.7	0.0041	*nd*
LB	328	AB-000224	MS-1795	12	102.2	100	>0.5	*nd*	NA	NA	45.8	3.6	63.0	>0.4	*nd*
LB	328	AB-000224	MS-1796	10	101.0	100	>0.5	*nd*	NA	NA	44.9	4.8	61.8	>0.2	*nd*
LB	328	AB-000224	MS-1797	1	101.7	100	>0.5	*nd*	NA	NA	44.7	3.3	61.5	>0.2	*nd*
LB	412	NA	AB-000317	NA	NA	100	NA	NA	NA	NA	54.2	4.5	100.0	NA	NA
LB	412	NA	AB-007088	3	94.1	100	>0.5	NA	NA	NA	46.1	1.3	84.9	>0.5	NA
LB	412	AB-000224	MS-1798	5	103.3	100	>0.5	*nd*	NA	NA	32.0	3.1	58.9	0.0891	*nd*
LB	412	AB-000224	MS-1799	15	105.4	100	>0.5	*nd*	NA	NA	32.5	4.4	60.0	>0.1	*nd*
LB	412	AB-007088	MS-1801	5	96.0	100	>0.5	>0.5	NA	NA	32.1	2.5	59.2	>0.1	>0.5
LB	412	AB-007088	MS-1802	2	79.7	100	>0.1	>0.5	NA	NA	23.9	2.4	44.0	<0.0001	0.0011
LB	412	AB-007088	MS-1803	6	90.9	100	>0.5	>0.5	NA	NA	28.7	4.2	52.9	0.0088	0.0748
LB	412	AB-007088	MS-1804	4	88.8	100	>0.5	>0.5	NA	NA	35.9	4.2	66.3	>0.8	>0.5
LB	412	AB-007088	MS-1805	1	87.8	100	>0.5	>0.5	NA	NA	32.2	4.5	59.4	>0.1	>0.5
LB	509	NA	AB-000317	NA	NA	100	NA	NA	NA	NA	66.2	7.3	100.0	NA	NA
LB	509	AB-000224	MS-1785	4	97.9	100	>0.5	*nd*	NA	NA	41.6	2.3	62.8	>0.1	*nd*
LB	509	AB-000224	MS-1788	7	98.1	100	>0.5	*nd*	NA	NA	41.0	3.7	61.9	0.0507	*nd*
LB	509	AB-000224	MS-1797	1	99.5	100	>0.5	*nd*	NA	NA	38.4	4.5	58.0	0.0044	*nd*
LB	510	NA	AB-000317	NA	NA	100	NA	NA	NA	NA	63.4	5.5	100.0	NA	NA
LB	510	AB-000224	MS-1790	4	101.2	100	>0.5	*nd*	NA	NA	45.3	2.3	71.4	>0.3	*nd*
LB	510	AB-000224	MS-1796	10	102.0	100	>0.5	*nd*	NA	NA	42.6	2.3	67.2	0.0309	*nd*
LB	510	AB-000224	MS-1798	5	102.0	100	>0.5	*nd*	NA	NA	40.6	2.8	64.1	0.0105	*nd*
LB	512	NA	AB-000317	NA	NA	100	NA	NA	NA	NA	66.8	6.1	100.0	NA	NA
LB	512	AB-000224	MS-1794	2	100.4	100	>0.5	*nd*	NA	NA	33.3	1.4	56.7	0.001	*nd*
LB	512	AB-000224	MS-1795	12	101.4	100	>0.5	*nd*	NA	NA	38.3	6.3	67.4	0.0508	*nd*
LB	512	AB-000224	MS-1799	15	101.6	100	>0.5	*nd*	NA	NA	39.1	3.8	68.7	>0.1	*nd*
LB	513	NA	AB-000317	NA	NA	100	NA	NA	NA	NA	59.5	6.3	100.0	NA	NA
LB	513	AB-000224	MS-1787	6	103.4	100	>0.5	*nd*	NA	NA	45.1	6.1	75.8	>0.2	*nd*
LB	513	AB-000224	MS-1792	8	100.7	100	>0.5	*nd*	NA	NA	42.6	4.7	71.6	0.0646	*nd*
LB	929	NA	AB-000317	NA	100.0	100	NA	NA	NA	NA	*nd*	*nd*	*nd*	NA	NA
LB	929	NA	AB-000224	16	101.6	100	>0.5	NA	NA	NA	*nd*	*nd*	*nd*	*nd*	NA
LB	929	AB-000224	MS-1797	1	103.5	100	>0.5	>0.5	NA	NA	*nd*	*nd*	*nd*	*nd*	*nd*
LB	509	NA	AB-000317	NA	NA	300	NA	NA	NA	NA	195.4	29.8	100.0	NA	NA
LB	509	AB-000224	MS-1785	4	100.0	300	>0.5	*nd*	NA	NA	120.4	10.5	61.6	0.0507	*nd*
LB	509	AB-000224	MS-1788	7	100.1	300	>0.5	*nd*	NA	NA	119.0	11.7	60.9	0.0449	*nd*
LB	509	AB-000224	MS-1797	1	100.0	300	>0.5	*nd*	NA	NA	142.0	20.1	72.7	>0.5	*nd*
LB	510	NA	AB-000317	NA	NA	300	NA	NA	NA	NA	170.6	11.8	100.0	NA	NA
LB	510	AB-000224	MS-1790	4	100.2	300	>0.5	*nd*	NA	NA	124.6	11.7	73.0	0.0538	*nd*
LB	510	AB-000224	MS-1796	10	100.2	300	>0.5	*nd*	NA	NA	130.4	16.9	76.4	>0.1	*nd*
LB	510	AB-000224	MS-1798	5	100.3	300	>0.5	*nd*	NA	NA	139.8	15.1	81.9	>0.4	*nd*
LB	512	NA	AB-000317	NA	NA	300	NA	NA	NA	NA	182.2	15.3	100.0	NA	NA
LB	512	AB-000224	MS-1794	2	100.0	300	>0.5	*nd*	NA	NA	111.2	10.3	61.0	0.0371	*nd*
LB	512	AB-000224	MS-1795	12	100.1	300	>0.5	*nd*	NA	NA	114.8	7.3	63.0	0.0763	*nd*
LB	512	AB-000224	MS-1799	15	100.0	300	>0.5	*nd*	NA	NA	126.2	15.1	69.3	>0.5	*nd*
LB	513	NA	AB-000317	NA	NA	300	NA	NA	NA	NA	214.4	10.2	100.0	NA	NA
LB	513	AB-000224	MS-1787	6	100.1	300	>0.5	*nd*	NA	NA	133.0	13.6	62.0	>0.1	*nd*
LB	513	AB-000224	MS-1792	8	100.3	300	>0.5	*nd*	NA	NA	128.0	15.4	59.7	0.0908	*nd*
BP	323	NA	AB-000317	NA		150	NA	NA	NA	NA	*nd*	*nd*	NA	NA	NA
BP	323	AB-000224	MS-1786	8		150	NA	NA	2.8	0.72, 11	*nd*	*nd*	NA	*nd*	*nd*
BP	419	NA	AB-000317	NA		150	NA	NA	NA	NA	98.7	15.3	100.0	NA	NA
BP	419	AB-000224	MS-1783	3		150	NA	NA	2.8	0.77, 10	67.2	10.7	68.1	0.001	*nd*
BP	419	AB-000224	MS-1785	4		150	NA	NA	1.1	0.27, 4.3	88.5	22.2	89.7	>0.5	*nd*
BP	419	AB-000224	MS-1790	4		150	NA	NA	1.7	0.45, 6.5	87.9	12.7	89.0	>0.5	*nd*
BP	419	AB-000224	MS-1797	1		150	NA	NA	1.2	0.30, 4.7	74.4	21.8	75.4	0.0553	*nd*
BP	422	NA	AB-000317	NA		150	NA	NA	NA	NA	141.1	25.0	100.0	NA	NA
BP	422	AB-000224	MS-1788	7		150	NA	NA	0.6	0.13, 2.6	113.4	35.0	80.3	0.0709	*nd*
BP	422	AB-000224	MS-1794	2		150	NA	NA	1.0	0.24, 4.3	131.4	24.7	93.1	>0.5	*nd*
BP	422	AB-000224	MS-1796	10		150	NA	NA	0.74	0.18, 3.1	120.2	12.9	85.2	>0.2	*nd*
BP	422	AB-000224	MS-1798	5		150	NA	NA	1.3	0.34, 5.2	119.2	29.6	84.5	>0.1	*nd*
BP	428	NA	AB-000317	NA		150	NA	NA	NA	NA	126.9	16.0	100.0	NA	NA
BP	428	AB-000224	MS-1787	6		150	NA	NA	0.70	0.11, 4.3	98.1	23.9	77.3	0.0788	*nd*
BP	428	AB-000224	MS-1792	8		150	NA	NA	6.5	1.6, 27	70.4	18.6	55.5	<0.0001	*nd*
BP	428	AB-000224	MS-1795	12		150	NA	NA	4.3	1.0, 19	96.8	31.9	76.3	0.0589	*nd*
BP	428	AB-000224	MS-1799	15		150	NA	NA	3.4	0.88, 13	96.8	18.8	76.2	0.0772	*nd*

LB, liver burden (n = 5 mice); BP, bite parasitemia (n = 10 mice); ID, identifier; NA, not applicable; HR, hazard ratio; CI, confidence interval; *nd*, not done

[a] mAb with greatest aggregate developability score from assay data is rank #1 among engineered variants of each, respective parental mAb (see methods); mAbs with equivalent scores have same rank number

[b] percent inhibition normalized to percent inhibition of AB-000317 tested at same dose in parallel

[c] *P* value from Dunn's multiple comparisons test, two-tailed

[d] two-tailed, logrank (95% CI) of the engineered variant compared to AB-000317

nature portfolio

Corresponding author(s): Katherine L Williams
Daniel E Emerling

Last updated by author(s): Oct 10, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

Plasmablasts were collected using BD FACSDiva software v8.0.3
SPR binding kinetic data was collected on the Carterra LSA using Epitope Software v1.5.
Next-gen sequences were collected using Roche GS FLX+ System Software v2.9
In vivo luminescence data was collected using Living Image v3.2 software (Perkin Elmer)
ELISA data were collected using Biotek Gen5 software v2 or Molecular Devices microplate reader with SoftMax-Pro GxP version 6.5.1

Data analysis

The kinetics titration data collected were first pre-processed in the NextGenKIT (Carterra) software v1.7. The data were then exported and analysed using the TitrationAnalysis tool (https://zenodo.org/record /7998652). Custom code for the tool available at https://github.com/DukeCHSI/TitrationAnalysis.

In vivo luminescence data were analyzed using Living Image v3.2 software (Perkin Elmer)

All statistical tests were performed using GraphPad Prism v8 or v9 except for the Bootstrap analyses and the Benjamini-Hochberg False Discovery Rate analysis which were performed using R v3.4.4 (cran.r-project.org).
Custom code used in R for bootstrap analyses is available at the site https://github.com/expositum/Mal71_CSPBinding_vs_Protection.
Custom code used in the TitrationAnalysis tool for SPR analyses is available at https://github.com/DukeCHSI/TitrationAnalysis

Cell counts for engineered production cell lines were analysed using Guava CytoSoft Data Acquisition and Analysis Software v3.2

ELISA data were analyzed using Biotek Gen5 software v2

IgG sequences were assembled to variable (V), diversity (D) and joining (J) gene segment assignments and somatic hypermutations (SHM) were identified using Somatic Diversification Analysis [SoDA, Volpe, J. M., Cowell, L. G. & Kepler, T. B. SoDA: implementation of a 3D alignment algorithm for inference of antigen receptor recombinations. Bioinformatics 22, 438–444 (2006)] and the IMGT human immunoglobulin germline database release, IMGT_202031
Structural models were made using Molecular Operating Environment, MOE, Molecular Operating Environment, 2022.02 Chemical Computing Group ULC

Other analyses were performed using Microsoft Excel (multiple versions)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

IgG sequence datasets Accession numbers for all paired heavy and light chain IgG sequences that were recombinantly expressed were provided through BankIt: 2749610: OR662637 - OR663656). The entire set of unique natively paired IgG sequences (n = 28,672) from PBs (n = 32,948) of RTS,S vaccines (n = 45) are available at the site, https://zenodo.org/record/8436761.

Requests for other datasets generated and/or analysed in the current study will be promptly reviewed by the corresponding authors (emerling@biosimplify.com or kwilliams@atreca.com) and a Material Transfer Agreement provided should the request be subject to intellectual property obligations. Materials subject to an MTA will be released pending execution. All other data/ materials not subject to an MTA will be provided within a reasonable timeframe following the initial request.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Previously reported in Regules, J. A. et al. Fractional third and fourth dose of RTS,S/AS01 malaria candidate vaccine: A phase 2a controlled human malaria parasite infection and immunogenicity study. J. Infect. Dis. 214, 762–771 (2016)] The ratio of females to males was 15:19 and 6:11 in the Fx017M, 012M groups respectively.
Population characteristics	Previously reported in Regules, J. A. et al. Fractional third and fourth dose of RTS,S/AS01 malaria candidate vaccine: A phase 2a controlled human malaria parasite infection and immunogenicity study. J. Infect. Dis. 214, 762–771 (2016)] Vaccinees included males and non-pregnant females aged 18–50 years who were free of any serious acute or chronic illness, as determined by clinical or physical examination, medical history records, or laboratory screening tests of hematologic, renal, and hepatic function; did not have a history of malaria; were seronegative for HBsAg, hepatitis C virus, and human immunodeficiency virus (HIV); and had the ability to comply with the study protocol. The mean age of subjects was 33.6 years. The ratio of females to males was 15:19 and 6:11 in the Fx017M, 012M groups respectively.
Recruitment	Previously reported in Regules, J. A. et al. Fractional third and fourth dose of RTS,S/AS01 malaria candidate vaccine: A phase 2a controlled human malaria parasite infection and immunogenicity study. J. Infect. Dis. 214, 762–771 (2016)] The target enrollment was 65 volunteers with consecutive (nonrandomized and open) allocation to study groups: the first 34 volunteers were to immunized according to a 0-, 1-, and 7-month schedule with a fractional third dose (Fx017M), the subsequent 17 volunteers were immunized with 3 full doses according to a 0-, 1-, and 2-month schedule (012M). For this study, we received PBMC samples from 30 of the 017M vaccinees and 15 of the 012M vaccinees.
Ethics oversight	The protocol was approved by the Walter Reed Army Institute of Research Institutional Review Board and the Western Institutional Review Board, and written informed consent was obtained from each subject before study procedures were initiated (ClinicalTrials.gov identifier: NCT01857869 and Regules, J. A. et al. Fractional third and fourth dose of RTS,S/AS01 malaria candidate vaccine: A phase 2a controlled human malaria parasite infection and immunogenicity study. J. Infect. Dis. 214, 762–771 (2016)]

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The main aim of this study was to downselect an antibody clone for potential development as a drug from sequence repertoires of RTS,S vaccinees. Sample size of vaccinees sequenced was predetermined based on the RTS,S clinical trial study protocol and the PBMC samples obtained from that trial. The sample size of plasmablasts collected for sequencing was based on the exhaustive use of the samples to obtain plasmablasts. The number of antibody lineages included in the first screen for discovery of CSP-binding antibodies was determined based on goal to analyze a majority of the most dominant lineages from all the vaccinees. The sample size of antibodies selected for functional tests was based on the limited throughput of animal model functional testing. Technical replicates used in the initial antigen binding screens were duplicates for each antibody, and were duplicates or triplicates for the . SPR binding experiments. In all these binding assays, standard replicate numbers were used as common in the field and no statistical determination was made for technical replicate numbers used. For SPR binding kinetics determination, the number of concentrations was determined by making a two-fold dilution series that provided enough data points to determine Koff, Kon, and KD values. For in vivo studies, each antibody was tested in 5 mice (liver burden model) or 10 mice (mosquito bite parasitemia model) as a sample size that minimizes the use of animals while still providing a means to statistically compare the activity of test articles to control conditions as reported elsewhere [Flores-Garcia, Y. et al. Optimization of an in vivo model to study immunity to Plasmodium falciparum pre-erythrocytic stages. Malar. J. 18, 426 (2019)]. Correlation analyses of sequence repertoire features, of binding data and protection status of vaccinees, of binding kinetics data and mutation (SHM) level data, of binding kinetics data and functional data, and of mutation (SHM) level data and functional data were all performed post-hoc with the available data that was obtained with the experimental goal of screening for, discovering and engineering an anti-CSP drug. Thus, no power analyses were performed to determine sample size for these correlation analyses prior to experimental design, however sample size would not appear limiting for those cases in which a statistically significant correlation or inverse-correlation was reported.
Data exclusions	No animals or data points were excluded from the analyses of animal model data. For SPR binding experiments involving the engineered variants, some kinetic data were excluded from KD calculations per predetermined, data quality, acceptance criteria: i) standard error of the estimated kon, koff and KD in each replicate ≤20% and ii) fold-change for all three parameters within the triplicate ≤3.
Replication	In cases where antibodies from vaccinees' repertoires were tested in CSP or peptide ELISA assays more than once, reproduced positive, negative, or indeterminate results (see Methods) were obtained for CSP, NANP6, or C-terminal binding. Results from other binding experiments are reported as mean values from all replicates. Reproducibility was based on technical replicates as described in the "Sample size" section above. For some in vivo functional assessments of particular antibodies, independent experiments were reproduced and outcomes of these reproductions are reported in the manuscript.
Randomization	Enrollment and participation in the RTS,S clinical trial was as previously reported in Regules, J. A. et al. Fractional third and fourth dose of RTS,S/AS01 malaria candidate vaccine: A phase 2a controlled human malaria parasite infection and immunogenicity study. J. Infect. Dis. 214, 762–771 (2016)]. The target enrollment was 65 volunteers with consecutive (nonrandomized and open) allocation to study groups: the first 34 volunteers were to immunized according to a 0-, 1-, and 7-month schedule with a fractional third dose (Fx017M), the subsequent 17 volunteers were immunized with 3 full doses according to a 0-, 1-, and 2-month schedule (012M). For this study, we received PBMC samples from 30 of the 017M vaccinees and 15 of the 012M vaccinees. Randomization of mice was not needed per protocols of the malaria models used. Randomization of PBMC samples and antibodies was not appropriate/relevant to the study
Blinding	Linkage of vaccine trial participant information (i.e. protection status and trial arm assignment) and PBMC samples, IgG sequences derived from the PBMC, and recombinant antibodies derived from the sequences was blinded to those researchers who isolated plasmablasts, sequenced IgG, and tested antibodies in initial binding and in vivo functional screens.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

anti-CD3-FITC (BioLegend, cat# 300440, clone UCHT1, 1:100), anti-CD14-FITC (Bio-Legend, cat# 325604, clone HCD14, 1:100), anti-CD19-BV421 (BioLegend, cat# 302234, clone HIB19, 1:100), anti-CD20-PE/cy7 (Biolegend, cat# 302312, clone 2H7, 1:100), anti-CD27-BV510 (BioLegend, cat# 302836, clone O323, 1:50), anti-CD38-A647 (BioLegend, cat# 303514, clone HIT2, 1:100), anti-IgA-FITC (Miltenyi, cat# 130- 113-475, clone IS11-8E10, 1:50), anti-IgM-FITC (BioLegend, cat# 314506, clone MHM-88, 1:50), anti-IgD-FITC (Biolegend #348206, clone IA6-2, 1:50)

Monoclonal antibodies discovered and characterized in the study are defined by their heavy and light chain variable regions' sequences and can be gene synthesized and recombinantly expressed from those sequences which will be provided via a public data repository (please see availability of data statement)

Validation

https://www.biolegend.com/en-us/products/fitc-anti-human-cd3-antibody-863?pdf=true&displayInline=true&leftRightMargin=15&topBottomMargin=15&filename=FITC%20anti-human%20CD3%20Antibody.pdf&v=20220902063018

https://www.biolegend.com/en-us/products/fitc-anti-human-cd14-antibody-3951?pdf=true&displayInline=true&leftRightMargin=15&topBottomMargin=15&filename=FITC%20anti-human%20CD14%20Antibody.pdf&v=20230114013553

https://www.biolegend.com/en-us/products/brilliant-violet-421-anti-human-cd19-antibody-7144?pdf=true&displayInline=true&leftRightMargin=15&topBottomMargin=15&filename=Brilliant%20Violet%20421%E2%84%A2%20anti-human%20CD19%20Antibody.pdf&v=20230114013553

https://www.bdbiosciences.com/content/bdb/paths/generate-tds-document.us.340955.pdf

https://www.biolegend.com/en-us/products/brilliant-violet-510-anti-human-cd27-antibody-8005?pdf=true&displayInline=true&leftRightMargin=15&topBottomMargin=15&filename=Brilliant%20Violet%20510%E2%84%A2%20anti-human%20CD27%20Antibody.pdf&v=20220831123135

https://www.biolegend.com/en-us/products/pe-cyanine7-anti-human-cd38-antibody-8420?pdf=true&displayInline=true&leftRightMargin=15&topBottomMargin=15&filename=PE/Cyanine7%20anti-human%20CD38%20Antibody.pdf&v=20230114013553

https://www.miltenyibiotec.com/US-en/products/iga-antibody-anti-human-is11-8e10.html#fitc:100-tests-in-200-ul

https://www.biolegend.com/en-us/products/apc-cyanine7-anti-human-igm-antibody-7403?pdf=true&displayInline=true&leftRightMargin=15&topBottomMargin=15&filename=APC/Cyanine7%20anti-human%20IgM%20Antibody.pdf&v=20220817065308

https://d1spbj2x7qk4bg.cloudfront.net/ja-jp/products/fitc-anti-human-igd-antibody-6683?pdf=true&displayInline=true&leftRightMargin=15&topBottomMargin=15&filename=FITC%20anti-human%20IgD%20Antibody.pdf&v=20230714033116

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	HEK293 cells (ATCC) CHO-K1 GS KO cells (Horizon Discovery)
Authentication	Authentication via certificate of authentication from supplier
Mycoplasma contamination	Production cell line, CHO-K1, was tested for contamination per FDA guidelines and requirements
Commonly misidentified lines (See ICLAC register)	*Name any commonly misidentified cell lines used in the study and provide a rationale for their use.*

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	Studies using mice were carried out using 6-8 weeks old C57BL/6 female mice (Charles River Labs), maintained at the animal facility of the Johns Hopkins Bloomberg School of Public Health. The assays using mice were performed in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The protocol was approved by the Animal Care and Use Committee of the Johns Hopkins University, protocol numbers, MO18H419 and MO21H417. No animals or data points were excluded from the analyses Approximately 3000 mice were used across the in vivo experiments reported in this study.
Wild animals	The study did not involve wild animals
Reporting on sex	All experiments were done using female mice. Our studies have shown that female mice are more susceptible to infection*. Thus, the infection assays have a dynamic range that allow us to differentiate better between antibodies with different protective efficacy. * Flores-Garcia et al., Malaria Journal volume 18, Article number: 426 (2019) Optimization of an in vivo model to study immunity to Plasmodium falciparum pre-erythrocytic stages
Field-collected samples	The study did not involve samples collected from the field
Ethics oversight	The assays using mice were performed in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The protocol was approved by the Animal Care and Use Committee of the Johns Hopkins University, protocols number MO18H419 and MO21H417

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	ClinicalTrials.gov identifier: NCT01857869 Human research participants were part of the phase 2a clinical trial of RTS,S/AS01B which has been reported previously [Regules, J. A. et al. Fractional third and fourth dose of RTS,S/AS01 malaria candidate vaccine: A phase 2a controlled human malaria parasite infection and immunogenicity study. J. Infect. Dis. 214, 762–771 (2016)]. The protocol was approved by the Walter Reed Army Institute of Research Institutional Review Board and the Western Institutional Review Board, and written informed consent was obtained from each subject before study procedures were initiated
Study protocol	Reported previously in Regules, J. A. et al. Fractional third and fourth dose of RTS,S/AS01 malaria candidate vaccine: A phase 2a controlled human malaria parasite infection and immunogenicity study. J. Infect. Dis. 214, 762–771 (2016)].
Data collection	Reported previously in Regules, J. A. et al. Fractional third and fourth dose of RTS,S/AS01 malaria candidate vaccine: A phase 2a controlled human malaria parasite infection and immunogenicity study. J. Infect. Dis. 214, 762–771 (2016)].
Outcomes	Reported previously in Regules, J. A. et al. Fractional third and fourth dose of RTS,S/AS01 malaria candidate vaccine: A phase 2a controlled human malaria parasite infection and immunogenicity study. J. Infect. Dis. 214, 762–771 (2016)].

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Plasmablast cells were isolated/collected by flow cytometry and used for sequencing of IgG. Gates are shown in Supplementary Fig. 9 with percentage of cells within each gate noted (statistics are not applicable) Plasmablast isolation, cloning, and sequencing were performed using BD FACSDiva soft-ware v8.0.3 and the previously published protocol78 with the following modifications. PBMC were stained with the following mAbs and dilutions: anti-CD3-FITC (BioLegend, cat# 300406, clone UCHT1, 1:100), anti-CD14-FITC (Bio-Legend, cat# 325604, clone HCD14, 1:100), anti-

	CD19-BV421 (BioLegend, cat# 302234, clone HIB19, 1:100), anti-CD20-PerCP/cy5.5 (BD, cat# 340955, clone L27, 1:15), anti-CD27-BV510 (BioLegend, cat# 302836, clone O323, 1:50), anti-CD38-PE/cy7 (BioLegend, cat# 356607, clone HB-7, 1:200), anti-IgA-FITC (Miltenyi, cat# 130- 113-475, clone IS11-8E10, 1:50), anti-IgM-APC/cy7 (BioLegend, cat# 314520, clone MHM-88, 1:100)
Instrument	BD FACSAria II
Software	BD FACSDiva software, v8.0.3
Cell population abundance	Purity of post-sort fractions could not be assessed in this study because IgG+ PBs were directly sorted one-cell-per-well into 96-well PCR plates containing hypotonic buffer that lysed the cells as intended for post-sort barcoding and sequening. Levels of target plasmablast cells among B cells and PBMC are reported in Supplementary Table 1
Gating strategy	Gating included both lymphoid and myeloid populations and excluded cell debris and aggregates in the FSC/SSC plot. Doublets and PI+ dead cells were gated out from the lymphoid and myeloid populations. Plasmablast target cells were then gated as the CD3-CD14-CD19+CD20–CD27+CD38++IgA–IgM– cells

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.