Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Nat Microbiol. Author manuscript; available in PMC 2021 July 15.

Published in final edited form as:

Nat Microbiol. 2021 March ; 6(3): 380–391. doi:10.1038/s41564-020-00858-9.

Structural basis for placental malaria mediated by *Plasmodium falciparum* VAR2CSA

Rui Ma[1], **Tengfei Lian**[2], **Rick Huang**[3], **Jonathan P. Renn**[4], **Jennifer D. Petersen**[5], **Joshua Zimmerberg**[5], **Patrick E. Duffy**[4,6], **Niraj H. Tolia**[1,*]

[1]Host-Pathogen Interactions and Structural Vaccinology Section, Laboratory of Malaria Immunology and Vaccinology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA

[2]Laboratory of Membrane Proteins and Structural Biology, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA

[3]Laboratory of Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

[4]Vaccine Development Unit, Laboratory of Malaria Immunology and Vaccinology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA

[5]Section on Integrative Biophysics, Division of Basic and Translational Biophysics, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA

[6]Pathogenesis and Immunity Section, Laboratory of Malaria Immunology and Vaccinology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA

Abstract

Plasmodium falciparum VAR2CSA binds to chondroitin sulfate A (CSA) on the surface of the syncytiotrophoblast during placental malaria. This interaction facilitates placental sequestration of malaria parasites resulting in severe health outcomes for both the mother and her offspring. Furthermore, CSA is presented by diverse cancer cells and specific targeting of cells by VAR2CSA may become a viable approach for cancer treatment. Here, we determined the Cryo-EM structures of the full-length ectodomain of VAR2CSA from *P. falciparum* strain NF54 in complex with CSA, and VAR2CSA from a second *P. falciparum* strain FCR3. The architecture of VAR2CSA is composed of a stable core flanked by a flexible arm. CSA traverses the core domain

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

*Correspondence: niraj.tolia@nih.gov (N.H.T.).

Author contributions

N.H.T. conceived the study, designed and supervised research. R.M. designed and carried out all experiments and data analysis. J.P.R with support from P.E.D., cloned the original expression plasmids to generate a panel of full-length VAR2CSA variants. T.L. assisted with grid freezing, Cryo-EM data collection and data processing of crosslinked VAR2CSA FCR3. R.H assisted with Cryo-EM data collection of VAR2CSA FCR3 and VAR2CSA NF54 CSA complex. J.D.P. and J.Z. help with the negative stain studies. R.M. and N.H.T. interpreted the data and wrote the manuscript, with input from all authors.

Competing interests

The authors declare no competing interests.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

by binding within two channels and CSA binding does not induce major conformational changes in VAR2CSA. The CSA-binding elements are conserved across VAR2CSA variants and are flanked by polymorphic segments, suggesting immune selection outside the CSA-binding sites. This work provides paths for developing interventions against placental malaria and cancer.

Women become more susceptible to malaria infection during pregnancy despite pre-existing immunity acquired from childhood, causing significant risk of severe outcomes for the mother and their offspring[1]. Placental malaria is caused by the accumulation of *Plasmodium falciparum*-infected erythrocytes in the placenta of pregnant women, resulting in high rates of maternal anemia, low birth weight, stillbirth and spontaneous pregnancy loss[1–3]. Each year, up to 200,000 infant deaths and 10,000 maternal deaths are attributed to malaria infection in pregnancy globally[4,5]. However, women naturally acquire resistance to placental malaria over successive pregnancies, providing a strong basis for the development of vaccines to prevent placental malaria[6,7].

P. falciparum expresses a family of proteins, referred to as erythrocyte membrane protein 1 (PfEMP1), that are translocated to the surface of the infected erythrocyte to enable adherence to different host organs and to evade the host immune response[8]. The leading placental malaria vaccine candidate VAR2CSA is a member of the PfEMP1 family that specifically binds to the syncytiotrophoblast surface receptor chondroitin sulfate A (CSA) leading to placental malaria[9,10]. Due to its large size [including an ~310 kDa extracellular domain (Fig. 1a)], production of VAR2CSA protein for vaccine development and scientific study has proven challenging[11]. Furthermore, the highly polymorphic nature of the extracellular domain of VAR2CSA in parasite isolates may hinder the development of a strain-transcending vaccine[12,13]. Lastly, vaccine-induced and naturally acquired immunity may differ in important ways that need to be carefully examined.

The ectodomain of VAR2CSA consists of an N-terminal sequence (NTS), six Duffy-binding like (DBL) domains, and three interdomain regions (ID) organized as shown in Fig. 1a. ID2 is also referred to as the cysteine-rich interdomain region ($CIDR_{PAM}$). Low resolution small-angle X-ray scattering and negative stain electron microscopy studies indicate the ectodomain of VAR2CSA adopts a compact structure[10,14,15]. In vitro studies suggest that the segment comprising DBL1X to ID2a is sufficient to bind CSA and is considered the minimal CSA-binding region, although DBL3X and DBL6ε have also shown CSA-binding activity and the full-length ectodomain exhibits higher affinity than any region alone[10,14,16–23]. Two candidate vaccines, PRIMVAC and PAMVAC derived from the VAR2CSA segments within DBL1X through ID2a are currently in phase I/II clinical trials[24–26] (Fig. 1a). Immunization with these VAR2CSA segments generates homologous inhibitory antibodies but limited heterologous activity against disparate strains[25,26]. These findings highlight the need to develop a strain-transcending vaccine against placental malaria based on VAR2CSA.

Intriguingly, diverse cancer cells express and present the form of chondroitin sulfate that is typically found exclusively in the placenta, and a recombinant VAR2CSA fragment (rVAR2) conjugated to therapeutics could inhibit tumor cell growth in vivo[27] (Fig. 1a). Thus, VAR2CSA has been used to develop platforms for cancer diagnosis and therapeutics[27–30].

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Despite the importance of VAR2CSA in both malaria and cancer, two diseases of global importance, critical information has been lacking about the specific recognition mechanism for VAR2CSA binding CSA. Here, we present the Cryo-EM structures of the full-length ectodomain of VAR2CSA in both ligand-binding and ligand-free states. The structures reveal that the CSA-binding sites reside in two binding channels within the core structure of VAR2CSA. This work elucidates the sequestration mechanism of placental malaria and has direct implications for development of malaria vaccines as well as cancer therapeutics and diagnostics.

Results

Overall structure of the CSA-VAR2CSA complex

We expressed VAR2CSA from parasite strain NF54 (VAR2CSA NF54) in Expi293 cells and purified it for the Cryo-EM study of VAR2CSA in complex with CSA (Extended Fig.1a,b). 6,196 movies were collected allowing for a 3.82 Å reconstruction of VAR2CSA NF54 in complex with CSA (Extended Data Fig. 2a). VAR2CSA NF54 exhibits an architecture comprised of a stable core and a flexible arm (Fig. 1a–c). Local refinement of the core improved the resolution to 3.36 Å, and local refinement of the arm resulted in a 4.88 Å map (Extended Data Fig. 2a–f). A CSA dodecamer spans the core domain and binds in a channel termed the major binding channel (Fig. 1b–d). Another potential binding site for CSA was observed in a second channel termed the minor binding channel with weak density that could be modeled as a CSA monosaccharide (Fig. 1b,d). This binding of CSA polymer within channels of VAR2CSA is reminiscent of the binding model proposed for EBA-175 binding to glycophorin A during *P. falciparum* invasion of erythrocytes, where the glycophorin A receptor feeds through channels created by EBA-175[31]. EBA-175 is a protein related to VAR2CSA that belongs to the erythrocyte binding-like (EBL) family involved in the recognition of sialic acid on erythrocyte glycoproteins during erythrocyte invasion by *P. falciparum*[32–34]. Our final model for VAR2CSA NF54 spans residues 32 to 2607 of VAR2CSA NF54 with a few flexible loops and ID1 omitted as these segments were not ordered in the reconstruction (Fig. 1c, Extended Data Fig. 2g–I and supplementary Table 1).

An interwoven domain architecture stabilizes VAR2CSA

VAR2CSA is primarily composed of α-helices and extensive loops that adopt an overall shape resembling the number 7 (Fig. 1c). CryoSPARC 3D variability analysis confirms that the region composed of DBL2X to ID3 forms a relatively stable core, while DBL5ε-DBL6ε forms a flexible arm and DBL1X exhibits some structural flexibility (Supplementary Video 1). The 6 individual DBL domains of VAR2CSA adopt the classical DBL domain fold, consisting of an α-helical core decorated by extensive loops[31,35–37] (Fig. 2a). The individual domains interact in an interwoven manner to stabilize the compact tertiary structure (Fig. 1c,d). DBL4ε, the most conserved DBL domain of the six[38], is located at the center of VAR2CSA and unites the whole structure by directly interacting with all the other domains except DBL1X and DBL5ε (Fig. 1c,d). DBL1X and DBL5ε are connected to DBL4ε via the NTS and ID3, respectively (Fig. 1c,d). The NTS (residue 32–49) is a twisted loop surrounding DBL1X and serves as the mortar holding together DBL1X and DBL4ε, with high conservation among diverse VAR2CSA strains (Fig. 2a–c). ID3 is a long helix that

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

closely interacts with ID2 and connects DBL5ε with the core (Fig. 2a,d). A total of 31 pairs of disulfide bridges were identified in the final model (Fig. 2e and supplementary Table 2).

Structural conservation within the EBL and PfEMP1 families

The VAR2CSA structure represents the first characterized structure of a full-length PfEMP1 protein, and provides the first structural models for DBL1X, DBL2X, ID2a, ID2b, ID3 and DBL5ε (Fig. 2a). We performed structural alignments for these domains using the DALI search[39]. As expected, DBL1X, DBL2X and DBL5ε adopt structures similar to other DBL domains from PfEMP1 and EBA-175[31,36,37,40] (Extended Data Fig. 3a–c). We also observed a tandem packing of the dual DBL domains DBL3X/4ε and DBL5ε/6ε. These dual DBL domains exhibit a twisted pattern reminiscent of other tandem packed DBL pairs of EBA-140 and EBA-175, although the angle between DBL domains differs (Extended Data Fig. 3d–g)[31,41]. EBA-140 and EBA-175 both belong to the EBL family that mediates the recognition of sialic acid on erythrocyte glycoproteins[32].

The arrangement of DBL2X-ID2 represents a conserved architecture within the PfEMP1 protein family. The structure of DBL2X-ID2 from VAR2CSA is similar to the DBL1α-CIDRγ domains of PfEMP1-VarO, although they adopt different DBL-ID/CIDR orientation (Extended Data Fig. 3h). VarO binds the ABO blood group trisaccharide that mediates rosetting of infected red blood cells[40]. The individual DBL domains (DBL2X and DBL1α) are structurally similar, and the VAR2CSA ID2b domain has a strong similarity to the VarO CIDRγ subdomain 2 despite low sequence similarity (Extended Data Fig. 3b,i and j). The DBL-ID/CIDR angle differs between VAR2CSA and PfEMP1-VarO, but this tandem arrangement suggests that the DBL-ID/CIDR pairing among other PfEMP1 family members may have a similar architecture. These structural delineations will better inform and define the diverse PfEMP1 domain architectures.

Multiple domains within the core domains create major and minor CSA-binding channels

Previous studies have shown that VAR2CSA tends to bind the sulfate-clustered domains of the chondroitin sulfate proteoglycans in the intervillous spaces of the placenta[42], and a minimum of a CSA dodecasaccharide is required for efficient binding[42–45]. Indeed, the atomic resolution reconstructions provided assignment of a CSA polymer comprising twelve monomers bound in a positively charged channel that is formed by NTS, DBL1X, DBL2X and DBL4ε. (Fig. 3a,b). We name this channel the major CSA binding channel. Five sulfated N-acetylgalactosamine-4-sulfate (ASG) and six glucuronic acid (BDP) residues could be unambiguously assigned and built into the density. Furthermore, density for an additional residue is observed at the start of the chain that can accommodate a monosaccharide, but this density was not of sufficient quality to facilitate adequate modeling of this single residue (Fig. 3a).

The major binding channel can be separated into two non-continuous CSA-binding sites (Fig. 3a). The first binding site (major binding site 1) is located on the surface of DBL2X and binds CSA residues BDP-8 to ASG-11 (Fig. 3a,c,d). The sulfate group of ASG-11 forms hydrogen bonds with N557 while BDP-10 has interactions with R829, K561 and the main chain of A822 (Fig. 3c). ASG-9 forms multiple hydrogen bonds with K562, N576, K828

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and Q832 (Fig. 3d). The interaction of CSA with major binding site 1 is further strengthened by the hydrogen bonds between BDP-8 and K828 (Fig. 3d).

The second binding site (major binding site 2) lies deep in the hole of the funnel-shaped channel and is surrounded by NTS, DBL1X, DBL2X and DBL4ε (Fig. 3a,b). Multiple hydrogen bonds are also formed in this region: ASG-5 with K835, E1880 and K1889; BDP-4 with K48; ASG-3 with K48 and R846; BDP-2 with R846 and the main chain of I1785 (Fig. 3e). Y45 further stabilizes the interaction by packing tightly with BDP-2 (Fig. 3e). BDP-6 and ASG-7 do not exhibit direct interactions with VAR2CSA and may serve to link the two binding sites together (Fig. 3f).

Weak density (the size of a single ASG monosaccharide) was also identified in a separate region of VAR2CSA, which we have termed the minor binding channel that forms a potential second binding site (Extended Data Fig. 4a). The minor binding channel is made up of the residues from the C-terminus of DBL2X and N-terminus of ID2a, two regions previously implicated in CSA binding[16]. Similar to the major binding channel, the minor channel is rich in positively charged residues (Extended Data Fig. 4b).

The CSA-binding residues in both channels are highly conserved among different VAR2CSA alleles (Fig. 3g and Extended Data Fig. 4c). In addition, while individual segments of VAR2CSA demonstrate CSA binding, the full-length protein binds CSA with far greater affinity than any segment alone[14,19,46]. The structure provides a clear rationale for these observations. DBL1X, DBL2X, ID2a, ID2b, DBL3X, DBL4ε and ID3 all interact extensively to create an interwoven architecture (Figure 1b,d). The CSA binding is likely dependent on an intact core structure implicating multiple domains in high affinity CSA binding.

VAR2CSA adopts preformed CSA-binding channels

In addition to the CSA-VAR2CSA complex, we also solved the structure of CSA-free VAR2CSA from the parasite strain FCR3 (Extended Data Figs 1 and 5 and Supplementary Table 1). The sequence of these two VAR2CSA alleles share a 79% identity (Fig. 4a). The structure of VAR2CSA FCR3 may potentially inform development of a strain-transcending vaccine by revealing any conformational changes due to CSA binding, as well as commonalities and differences between strains. Additionally, the current placental malaria candidate vaccines are based on sequences from VAR2CSA NF54 and VAR2CSA FCR3, meaning that an FCR3 structure would facilitate comparison with existing candidate vaccines[25,26].

We determined the Cryo-EM structure of the CSA-free full-length VAR2CSA FCR3 to a resolution of 4 Å after collecting 10,108 movies (Extended Data Fig. 5a–c). The reconstructed map of apo VAR2CSA FCR3 exhibits a similar shape to the CSA bound VAR2CSA, and also resembles the number 7 with a stable core and flexible arm (Supplementary Video S2). Local refinement of the arm resulted in a 4.7 Å map (Extended Data Fig. 5d). In order to further improve the resolution and the accuracy of our atomic model, we cross-linked the full-length ectodomain under mild conditions and collected a second dataset of 4,739 micrograph movies. This dataset resulted in a reconstruction of the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

stable core comprising DBL1X to ID3 to 3.4 Å resolution, enabling accurate model building for this segment that comprises the core of VAR2CSA (Fig. 4b; Extended Data Figs. 5e–g, 6 and Supplementary Table 1). The 4.7 Å reconstruction from masked local refinement of DBL5ε and 6ε allowed docking and refinement of the C-alpha positions of DBL5ε as well as the available crystal structure of DBL6ε (PDB:2Y8D) into this map (Extended Data Fig. 6g)[47]. Our final model for VAR2CSA FCR3 spans residues 23 to 2602 of VAR2CSA with a few flexible loops and ID1 omitted, as these segments were not ordered in the reconstruction (Extended Data Fig. 6d). Comparison of the DBL1X-ID3 map generated from the crosslinked and non-crosslinked sample reveals no noticeable conformational changes in the core, indicating the crosslink did not affect conformation (Fig. 4c).

No major conformation changes were observed between the structures of CSA-bound and CSA-free VAR2CSA (Fig. 4c). The structural similarity between VAR2CSA FCR3 and NF54 also suggests different VAR2CSA variants are likely to have similar overall architecture (Fig. 4c). CSA could be well docked in the corresponding major and minor binding channels on VAR2CSA FCR3, which is similarly positively charged (Fig. 4c,d). This suggests that the CSA-binding mode we identified is conserved between strains, and that VAR2CSA does not require major conformational changes to enable CSA binding. However, some flexibility is observed in the region DBL1X-DBL2X, suggesting limited flexing of the molecule may facilitate CSA binding (Supplementary Video 2).

Analysis of VAR2CSA variability and placental malaria candidate vaccine designs

High sequence polymorphism among diverse VAR2CSA variants is one of the major barriers to strain-transcending vaccine development[38,48]. We analyzed the conservation of fourteen VAR2CSA sequences and mapped this onto the structure[49] (Fig. 5a). Residues in the CSA-binding sites within the major and minor binding channels are conserved, but the flanking regions are not (Fig. 5b). The high conservation of the residues within both channels that directly bind CSA indicate these residues are under selective pressure to be maintained across strains. These results suggest that all strains retain these residues to ensure CSA binding. The variability observed in the flanking regions that are distant from the CSA-binding residues and do not directly contact CSA suggests that variability at these positions should not impact CSA binding, but may play a role in immune evasion.

The CSA-binding site 2 is buried deeply in the major binding channel and may not be accessible to antibodies (Fig. 5b). Although binding site 1 is exposed on the VAR2CSA surface, the DBL2X surface surrounding the conserved CSA binding residues is highly heterogeneous among diverse VAR2CSA strains (Fig. 5b). Moreover, there is also extensive polymorphism surrounding the conserved residues within the minor CSA binding channel (Fig. 5b). This heterogeneity likely reflects variation induced under host immune pressure. Other than the key CSA-binding residues, a large number of the surface residues are polymorphic among different VAR2CSA strains (Fig. 5a).

The interwoven domain architecture identified in the structure is consistent with the finding that multiple domains play a role in binding CSA, as multiple domains create the binding channels. PAMVAC and PRIMVAC both include DBL2X, and this domain will help to generate CSA-blocking antibodies as it is the major domain contributing residues to binding

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

site 1. However, sequence variability surrounding binding site 1 (Fig. 5b) could potentially limit the induction of strain-transcending antibodies. Indeed, both candidate vaccines demonstrated low heterologous inhibitory activity[25,26]. PAMVAC and PRIMVAC contain only a portion of the major binding channel and this may explain the limited protection data (Fig. 5c,d). The structure of full-length VAR2CSA reveals larger CSA-binding sites with conserved targets for strain-transcending antibodies. This information will guide improvements on existing candidate vaccines and facilitate structure-based design of a strain-transcending placental malaria vaccine.

Epitopes mapping on VAR2CSA

The structure of full-length VAR2CSA provided a template to investigate previously discovered antibody epitopes. We mapped known epitopes on the structure (Fig. 6). Four multigravidae sera with inhibitory activity showed enhanced binding to distinct linear peptides using overlapping peptide scanning of DBL4ε[50]. All the sera showed antibody binding to peptides P23-P25 and one sample also showed reactivity to peptides P45 and P57. Interestingly, mapping of these peptides on the 3D structure revealed all these peptides cluster together and are located at the entrance of the deeply buried binding site 2 of the major CSA binding channel (Fig. 6). Separately, naturally acquired antibodies to ID1-DBL2-ID2a and DBL4ε recombinant constructs were found to have inhibitory activity against both homologous and heterologous isolates[51], and these results are consistent with the structural analysis identifying these domains as important for CSA binding. We mapped other known epitopes of antibodies from multigravid women (Fig.6). The epitopes of PAM8.1, which is an antibody derived from multigravid woman, was mapped to a strain-specific loop region on DBL3X[52]. However, this loop is not visible in the structure (Fig. 6). Peptide P62 found within DBL3X and peptide P63 within DBL5ε are two peptides that react strongly with Tanzanian female plasma[53]. Lastly, peptides P20 and P23 are two cryptic epitopes on DBL5ε which are shown to cross-react with the antibodies derived from *Plasmodium vivax* DBP (PvDBP)[54]. However, whether these peptide epitopes are neutralizing epitopes, and the inhibitory mechanisms of these antibodies require further study.

Discussion

The ability to sequester in different organs, combined with sophisticated antigenic diversity, has made *P. falciparum* the deadliest malaria species that infects humans[8]. Malaria during pregnancy is a major problem in sub-Saharan Africa, affecting an estimated 150 million pregnant women annually[1]. Women can become susceptible to malaria infection during pregnancy despite the immunity that might have developed from prior *P. falciparum* infections. Pregnant women may also serve as a reservoir for parasites which poses challenges to malaria eradication[1]. As parasites continue to develop drug resistance and new drugs entail potential teratogenesis, an effective vaccine to prevent placental malaria is urgently needed[1,55].

The Cryo-EM structure of VAR2CSA in CSA-bound and CSA-free states determined here support a model of binding depicted in Fig. 6. We identified a major CSA-binding channel

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

that has two non-continuous CSA-binding sites, and a potential minor CSA-binding channel on VAR2CSA which are pre-formed by multiple domains (Supplementary video 3). Although most of the CSA binding residues are highly conserved among various VAR2CSA alleles, a few residues at the openings of the binding sites exhibit slight polymorphism (Fig. 3g). In addition, the conserved residues are flanked by highly polymorphic residues (Fig. 5b). These variabilities may contribute to diverse binding affinity and PM severity of various VAR2CSA isolates[56]. The surface-exposed binding site 1 of the major binding channel is formed solely by DBL2X (Fig. 3a). The buried binding site 2 of the major binding channel and the minor binding channel are formed by the NTS, DBL1X, DBL2X, ID2a and DBL4ε domains (Fig. 3a). The finding that DBL2X appears in all CSA-binding sites suggests its central role in CSA binding. This is consistent with prior studies that identified DBL2X as central to the minimal CSA-binding region suggested for VAR2CSA[10,16], and the fact that DBL2X is included in both of the two candidate vaccines currently in clinical trials for placental malaria[25,26] (Fig. 1a). However, the multidomain binding model identifies all the CSA-binding regions and explains why the full-length VAR2CSA has much stronger CSA-binding affinity than any individual or short continuous domains[14] (Extended Data Fig. 7a). This study also identified DBL4ε as a key component of the CSA-binding channel. The binding residues of DBL4ε are buried in the hole of the channel and they work together with segments from the NTS, DBL1X and DBL2X to form the binding site 2 of the major binding channel.

The similar overall architecture of VAR2CSA from parasite strains NF54 and FCR3 implies that the VAR2CSA adopts a conserved shape. Some VAR2CSA proteins have been shown to have an additional DBL domain termed DBL7ε[57]. This DBL domain would be connected to the C-terminus of DBL6ε that is fully solvent exposed and away from both the arm and core of all structures reported here. DBL7ε can readily be accommodated as an extension of the arm and is unlikely to alter the remaining architecture of VAR2CSA.

One caveat of this study is we used CSA from bovine trachea which consists of a mixture of CSA with different sulfation patterns and of different lengths. Although five of the CSA disaccharides are fully sulfated in the structure, we cannot determine the sulfation status of the first CSA monosaccharide. Since the CSA completely traverses through the binding channel of VAR2CSA, it is also plausible that VAR2CSA may slide along a CSA chain to search for a highly sulfated cluster prior to strong binding. Furthermore, the fact that CSA is tethered to the proteoglycan in the placenta might facilitate the binding of multiple CSA glycans to the different CSA binding channels on one VAR2CSA molecule or distinct VAR2CSA molecules that are located on a same knob[58,59]. A second caveat is that while we observe density at two locations described as the major and minor binding channels further studies are required to establish the relative importance of each channel in binding. Recent study suggests that phosphorylation of VAR2CSA at residues S429, S433 and T934 is associated with enhanced adhesive properties[60]. S429 and S433 are located on ID1 which is flexible and disordered in the final reconstruction of this study, and T934 does not directly mediate the CSA binding. Therefore, how these three phosphoresidues impact the adhesion requires further investigation. Finally, while the resolution of the majority of the core is at atomic resolution the resolution for the arm is lower and should be interpreted accordingly.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The high variability of VAR2CSA from distinct *P. falciparum* strains poses a challenge to the development of strain-transcending vaccines for placental malaria. Mapping the VAR2CSA sequence variability onto the 3D structure of VAR2CSA shows the CSA-binding site 1 on DBL2X is highly conserved but is surrounded by highly polymorphic residues (Fig. 5b). This explains the low heterogenous inhibitory activity observed for the placental malaria candidate vaccines PRIMVAC and PAMVAC[25,26]. The highly polymorphic segments likely impact antibodies that bind at or close to the CSA-binding sites, preventing the development of antibodies capable of binding to the VAR2CSA variants. The structure of VAR2CSA bound to CSA presented here serves as a template to design and develop vaccines against placental malaria that will overcome strain-specific responses by focusing the immune response to conserved regions.

Multiple pieces of evidence suggest the immunogens that encompass the region NTS-DBL2X can bind to antibodies from multigravid women living in pandemic regions, and can induce protective antibodies in clinical trials[25,26,51,61]. Intriguingly, the previously identified linear peptide epitopes on DBL4ε reside right next the major CSA binding channel. It is possible that these linear peptides may be part of larger conformational epitopes that target the major CSA binding channel.

The limited structural information of full length PfEMP1 proteins has hampered progress towards understanding PfEMP1 host-parasite interactions and vaccine development. Low resolution structures of two other PfEMP1 proteins solved by Cryo-electron tomography and SAXS suggests they adopt shapes that mimic either a crescent or a boomerang[62–64]. However, a majority of PfEMP1 proteins utilize the N-terminal DBL and CIDR domains that correspond to the regions surrounding DBL2X of VAR2CSA to bind diverse host-receptors[65]. The atomic resolution structural information of how these segments in VAR2CSA bind CSA serves as a framework to understand PfEMP1 binding to diverse receptors. Together, these results suggest different PfEMP1 proteins may adopt various 3D structures, but they may utilize a conserved N terminal structure for receptor binding.

Interestingly, the form of CSA bound by VAR2CSA is exclusively expressed in the placenta in healthy individuals, but is expressed and presented in cells from diverse cancers of epithelial and mesenchymal origin[27]. This expression allows for the specific targeting of cancer cells by delivering therapeutics that utilize VAR2CSA as a carrier, and for VAR2CSA-based cancer diagnostics[28,29]. Clear structural definition of the functional segments from VAR2CSA required to bind CSA will lead to improvements for placental malaria vaccine development as well as cancer therapeutics and diagnostics (Extended Data Fig. 7b). rVAR2, which is composed of the DBL1X to ID2a domains of VAR2CSA, has been shown to specifically recognize cancer cells and can be conjugated with drugs to inhibit tumor growth[27]. rVAR2 comprises a similar region as the sequences used in PAMVAC and PRIMVAC (Fig. 1a). However, rVAR2 lacks the critical elements for full CSA binding provided by the NTS and DBL4ε that form the complete CSA-binding channel (Extended data Fig.8). Improving the affinity of VAR2CSA fragments for cancer therapy by structure-guided design may allow for improved treatments that require lower doses for efficacy.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

In summary, this study of VAR2CSA rationalizes available antibody-binding and receptor-binding observations and defines the CSA-binding elements that comprise conserved segments of VAR2CSA to target for strain-transcending protective immunity. This information will support precise design of vaccines to provide much needed medical countermeasures against placental malaria and will inform the development of potent targeted cancer therapeutics and diagnostics.

Methods

Purification of VAR2CSA NF54 and FCR3 in Expi293 cells

The wild type VAR2CSA NF54 and VAR2CSA FCR3 were expressed in Expi293 (Thermo Fisher) cells according to the manufacturer's protocols. In brief, the cells were grown shaking at 37°C and 8% CO_2, maintaining cultures at continuous log phase growth (3.0–5×10^6) for 3–4 passages after thawing. The day before transfection, 500 mL of culture was seeded at a density of 2.5–3×10^6 cells/mL in a 2 L flask. The day of transfection, cells were diluted back to 2.5–3×10^6 prior to transfection. The plasmid DNA was diluted with 25 mL of Opti-MEM I medium (Thermo Fisher) to a final concentration of 1μg/mL.

Then,1.4ml ExpiFectamin™ 293 Reagent (Thermo Fisher) was diluted with 25 ml Opti-MEM I medium, gently mixed and incubated at room temperature for 5 minutes. The diluted ExpiFectamine™ 293 Reagent was then added to the diluted plasmid DNA, mixed by swirling, and incubated at room temperature for 20 minutes. The mixture was added to the cells slowly while swirling the flask. The flask was returned to the incubator at 37°C and 8% CO_2. After 20 hours of incubation, ExpiFectamine™ 293 Transfection Enhancer 1 (Thermo Fisher) and ExpiFectamine™ 293 Transfection Enhancer 2 (Thermo Fisher) were added to the transfection flask.

Purification of VAR2CSA NF54 and FCR3

The cultures were centrifuged at 5000 rpm for 15 min five days post-transfection. The supernatant was collected and loaded onto Ni Sepharose™ Excel columns (GE Healthcare), which were manually packed in a glass gravity column. The column was washed twice with 10 column volumes of wash buffer (25mM HEPES, pH 7.4, 150mM NaCl, 25mM imidazole) and eluted with 5 column volumes of elution buffer (25mM Tris-HCl, pH 7.4, 150mM NaCl, 250mM imidazole). The elutes were concentrated with a 100kDa cutoff centrifugal filter unit (Millipore Sigma) to 1 ml and further purified by size-exclusion chromatography (Superose 6 Increase 10/300, GE Healthcare) in buffer A (10mM HEPES, pH 7.4, 100mM NaCl). The peak fractions were collected and verified by SDS-PAGE before EM grids preparation.

On-column crosslinking of VAR2CSA FCR3

In order to mildly stabilize the protein, on-column crosslinking of VAR2CSA FCR3 was performed as described[66]. First, a bolus of glutaraldehyde (200μl 0.25% v/v) was injected to a pre-equilibrated Superose 6 Increase 10/300 column in buffer A and run at 0.25 ml/min for 16 min (a total of 4 ml buffer). Then, the column flow was paused, and the injection loop was flushed using buffer followed by injection of purified VAR2CSA FCR3 (200 μl volume,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

at 3 μM concentration). Subsequently, the column was run at 0.25 ml/min and 0.3 ml fractions were collected for EM grids preparation.

VAR2CSA-CSA complex reconstruction

The Chondroitin sulfate A sodium salt from bovine trachea (Sigma) was dissolved in buffer A to 10 mg/ml. Then VAR2CSA was mixed with CSA at a molar ratio of 1:4. The mixture was incubated on ice for 30 min before EM grids preparation.

Cryo-EM grid preparation and data collection

The homogeneity of samples was first assessed by negative-stain EM with 0.7% (w/v) uranyl formate or 1% uranyl acetate as described[67]. Before preparing grids for Cryo-EM, the freshly purified protein sample was centrifuged at 13,000 g for 2 min to remove potential protein aggregates, and the protein concentration was measured with a NanoDrop spectrophotometer (Thermos Fisher Scientific). The final protein concentration used for Cryo-EM grid preparation is 0.8 mg/ml.

The protein sample was kept on ice before grid preparation. A 3.5 μL aliquot of protein was applied to a glow-discharged Quantifoil 300 mesh 1.2/1.3 carbon grid (Quantifoil) that had been glow-discharged for 90s at 10 mA with PELCO easiGlow Glow Discharge Set. VAR2CSA FCR3 and the VAR2CSA FCR3 crosslinked samples were blotted for 3s and VAR2CSA NF54 CSA complex was blotted for 2s with a blot force of 3 and 55/20mm filter paper (TED PELLA) before plunged into liquid ethane with a Vitrobot Mark VI (FEI) set at 16 °C and 100% humidity. After screening multiple grids, three grids made with the samples VAR2CSA NF54 in complex with CSA, VAR2CSA FCR3 alone and the crosslinked VAR2CSA FCR3 were chosen for data collection based on the evaluation of data quality.

The NF54+CSA and FCR3 dataset were collected on the 300keV Titan Krios with Gatan BioQuantum Image Filter in NIH National Cancer Institute (NCI)/NICE facility. The images were recorded with a 20 eV slit post-GIF K2 Summit camera in super-resolution counting mode at a nominal magnification of 130,000× and a defocus range from −0.7 to −2.0 μm. Exposures of 8s were dose-fractionated into 40 frames (200 ms per frame), with an exposure rate of 8 electrons·pixel^{-1} ·s^{-1}, resulting in a total exposure of 57 electrons·Å^{-2}, The data collection was automated using the SerialEM software package[68].

The FCR3 crosslink dataset was collected on Titan Krios electron microscopes in NIH Multi-Institute Cryo-EM Facility (MICEF). The images were recorded with a K2 Summit camera equipped with a Gatan Quantum LS imaging energy filter with the slit width set at 20 eV in counting mode at a nominal magnification of 130,000× and a defocus range from −1.0 to −2.0 μm. Exposures of 10s were dose-fractionated into 50 frames (200 ms per frame), with an exposure rate of 71.2 electrons·Å^{-2}, The data collection was automated using the Leginon software package[69].

Image processing

6,196 dose-fractionated movies of VAR2CSA NF54 + CSA were collected. The processing was done within cryoSPARC (v2.14.2)[70]. Motion correction was done by cryoSPARC's

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Patch motion correction with an output F-crop factor of ½. CTF estimation for each micrograph was calculated with Patch CTF estimation. Particles were autopicked from each micrograph with the blob picker from cryoSPARC and then sorted by 2D classification for two rounds to exclude bad particles. 858,299 particles were selected. The particles were used to generate an ab initio map in cryoSPARC. Particles were classified into 5 classes using the low-pass-filtered (30 Å) ab initio map as a template. Class 1 and Class 4 with a total of 299,571 particles, which has a clear map of the core region, was selected to conduct NU-refinement and generated a 3.5 Å map. A mask covering core regions were then used to perform local refinement and generated a 3.36 Å map. The map of the core is local filtered with a b-factor of −76.4 in Fig. 1c. Class 1 which has a clear density of the whole protein were selected solely with 157,702 particles to perform NU-Refinement and generated a 3.87 Å map of the full length complex. A mask covering DBL5ε and 6ε regions were then used to perform local refinement and generated a 4.88 Å map.

100,108 dose-fractionated movies of VAR2CSA FCR3 were collected on a 300-kV Titan Krios (FEI) equipped with a K2 Summit direct electron detector (Gatan). Similarly, the processing is done within cryoSPARC (v2.14.2)[70]. Motion correction was done by cryoSPARC's Patch motion correction with an output F-crop factor of ½. CTF estimation for each micrograph was calculated with Gctf (v1.06, https://www2.mrc-lmb.cam.ac.uk/research/locally-developed-software/zhang-software/)[71]. Particles were autopicked from each micrograph with the blob picker from cryoSPARC and then sorted by 2D classification for two rounds to exclude bad particles. 783,088 particles were selected. The dataset contained 783,088 particles. The particles were used to generate an ab initio map in cryoSPARC. Particles were classified into 10 classes using the low-pass-filtered (30 Å) ab initio map as a template. Class 4 with a total of 271,442 particles was selected to conduct NU-refinement and generated a 4 Å map. A mask covering DBL5ε and DBL6ε domains were then used to perform local refinement and generated a 4.69 Å map

4,739 dose-fractionated movies of VAR2CSA FCR3 were collected. The processing was also done within cryoSPARC[70]. Full frame motion correction was done by cryoSPARC's own implementation. CTF estimation for each micrograph was calculated with Gctf. 2,010,465 articles were autopicked from each micrograph with the blob picker from cryoSPARC and then sorted by 2D classification for two rounds to exclude bad particles. 505,409 particles were selected. The particles were used to generate an ab initio map in cryoSPARC. Particles were classified into 3 classes using the low-pass-filtered (30 Å) ab initio map as a template. Class 1 with a total of 319,520 particles was selected to conduct NU-refinement and generated a 3.52 Å map. A mask covering the core regions were then used to perform local refinement and generated a 3.38 Å map.

Model building and refinement

We first built up the model for the core of VAR2CSA FCR3 crosslinked. The crystal structures of DBL3X+4ε (PDB ID 4P1T)[46] was used as a starting model and was fitted and refined into the Cryo-EM density map with PHENIX (v1.18.2) 'Dock in map' and 'Real-space refinement[72]. The successful docking and the clear fitting of the DBL3X+4 side chains with the density indicated the fitting was correct. The clear density of a α-helix (ID3)

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

that connects the C-terminus of DBL4ε with the flexible arm which has a density of two tandem DBL domains helped us confirm the core is made up of DBL1X to DBL4ε while the arm is consists of DBL5ε and 6ε. The structures of DBL1X, DBL2X and ID2b were predicted from Phyre2 (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index)[73] and then fitted and refined into the map by PHENIX (v1.18.2). The missing regions were manually built in COOT (v0.9)[74]. The atomic model for the core was refined using phenix.real_space_refine global minimization (default), morphing and simulated annealing rama potential[75].

The model of the VAR2CSA crosslink core was used to build the VAR2CSA FCR3 structure by docking the model into the VAR2CSA FCR3 map and auto refined by PHENIX. To build the arm region of VAR2CSA, we used the crystal structure of DBL6ε (PDB ID 2Y8D)[47] and a predicted DBL5ε structure with Phyre2. The structures were fit in the Cryo-EM density map from local refinement with Chimera using the "fit in map" tool[76]. The atomic model for the arm was refined using phenix.real_space_refine global minimization (default), morphing and simulated annealing rama potential.

The model of core and arm region of CSA-VAR2CSA NF54 complex was built separately by fitting the corresponding model of VAR2CSA FCR3 into the map and manually mutating the residues and fragments adjustment. The CSA model was built with the C4S tetrasaccharide from the structure of the Shh–Chondroitin-4-Sulfate (C4S) Complex (PDB ID 4C4M)[77]. The atomic models was refined using phenix.real_space_refine global minimization (default), morphing and simulated annealing rama potential. Then we combined the models of the core and arm by fitting both of the maps together.

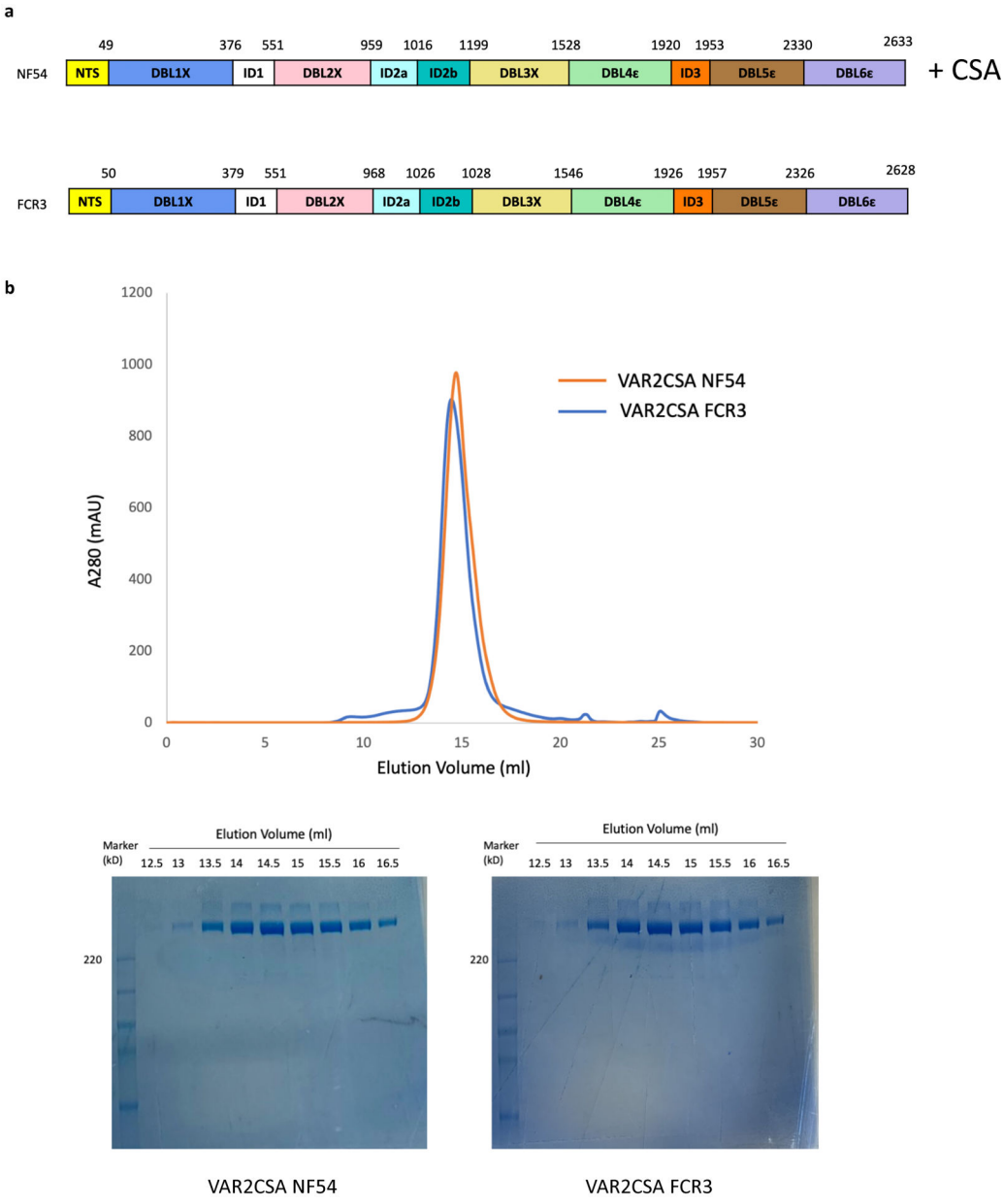
Structural and map figures were prepared in Chimera (v1.13.1, https://www.cgl.ucsf.edu/chimera/)[76], ChimeraX (v1.0, https://www.rbvi.ucsf.edu/chimerax/)[78] which are developed by UCSF, and PyMOL (v2.1, https://pymol.org/2/).

Data availability

Source data are provided with this paper, and the data that support the findings of this study are available from the corresponding author upon request. Atomic coordinates have been deposited at the Protein Data Bank under accession numbers 7JGD (VAR2CSA core crosslink), 7JGE (VAR2CSA FCR3 core), 7JGF (VAR2CSA FCR3 DBL5 and DBL6), 7JGG (VAR2CSA NF54 DBL5 and DBL6), and 7JGH (VAR2CSA NF54 + CSA core), and cryo-EM density maps have been deposited at the Electron Microscopy Data Bank under accession numbers EMD-22323 (VAR2CSA core crosslink), EMD-22324 (VAR2CSA FCR3), EMD-22325 (VAR2CSA FCR3 DBL5 and DBL6), EMD-22326 (VAR2CSA NF54 DBL5 and DBL6), and EMD-22327 (VAR2CSA NF54 core + CSA).

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

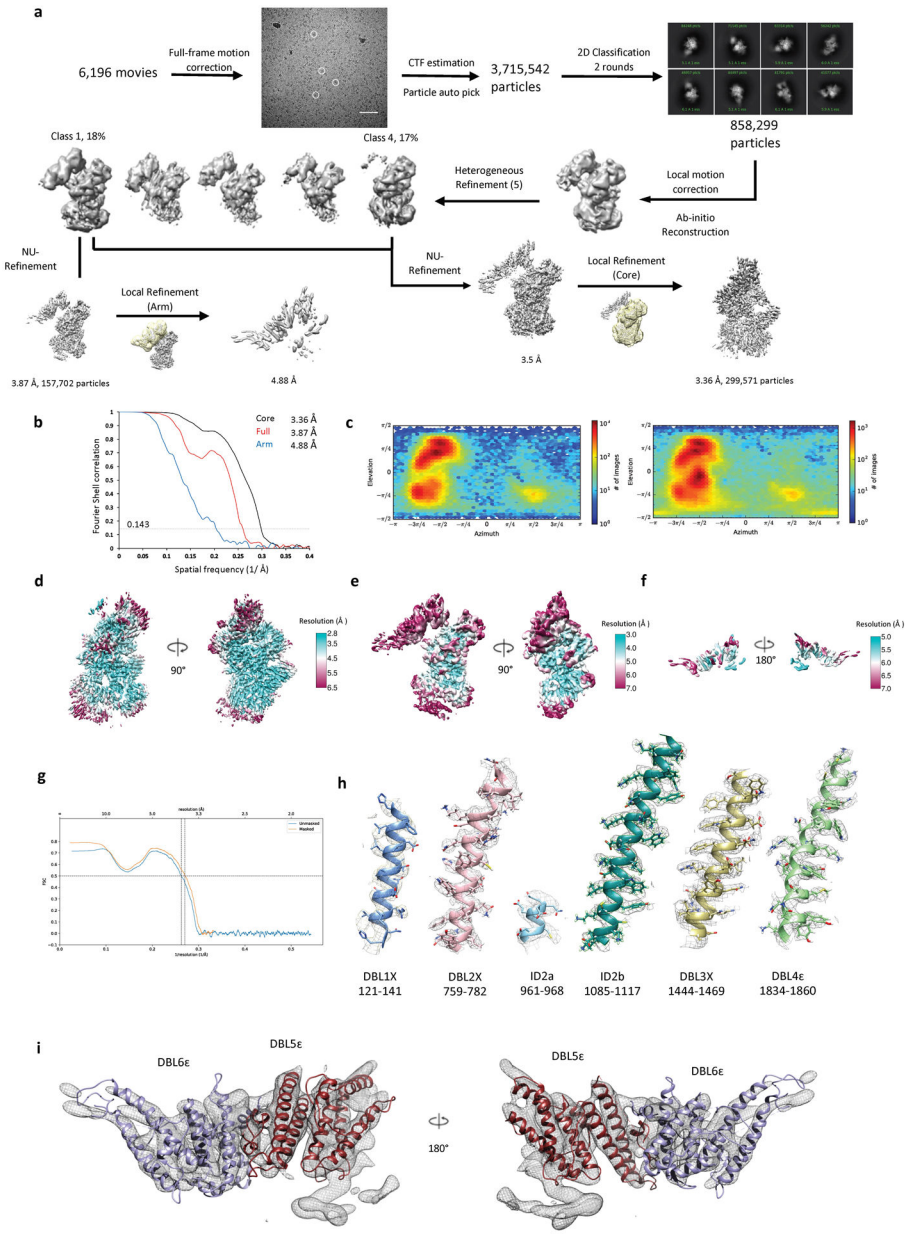
Extended Data



Extended Data Fig. 1. VAR2CSA Protein purification.

a, Domain boundaries of VAR2CSA NF54 and VAR2CSA FCR3 ectodomains we used in the structural analysis. b, Top: Size Exclusion Chromatography (SEC) profile of the VAR2CSA NF54 (orange) and VAR2CSA FCR3 proteins. Bottom: SDS PAGE analysis of the corresponding SEC fractions of VAR2CSA NF54 (left) and VAR2CSA FCR3 (right). Similar data were obtained from three independent purifications.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



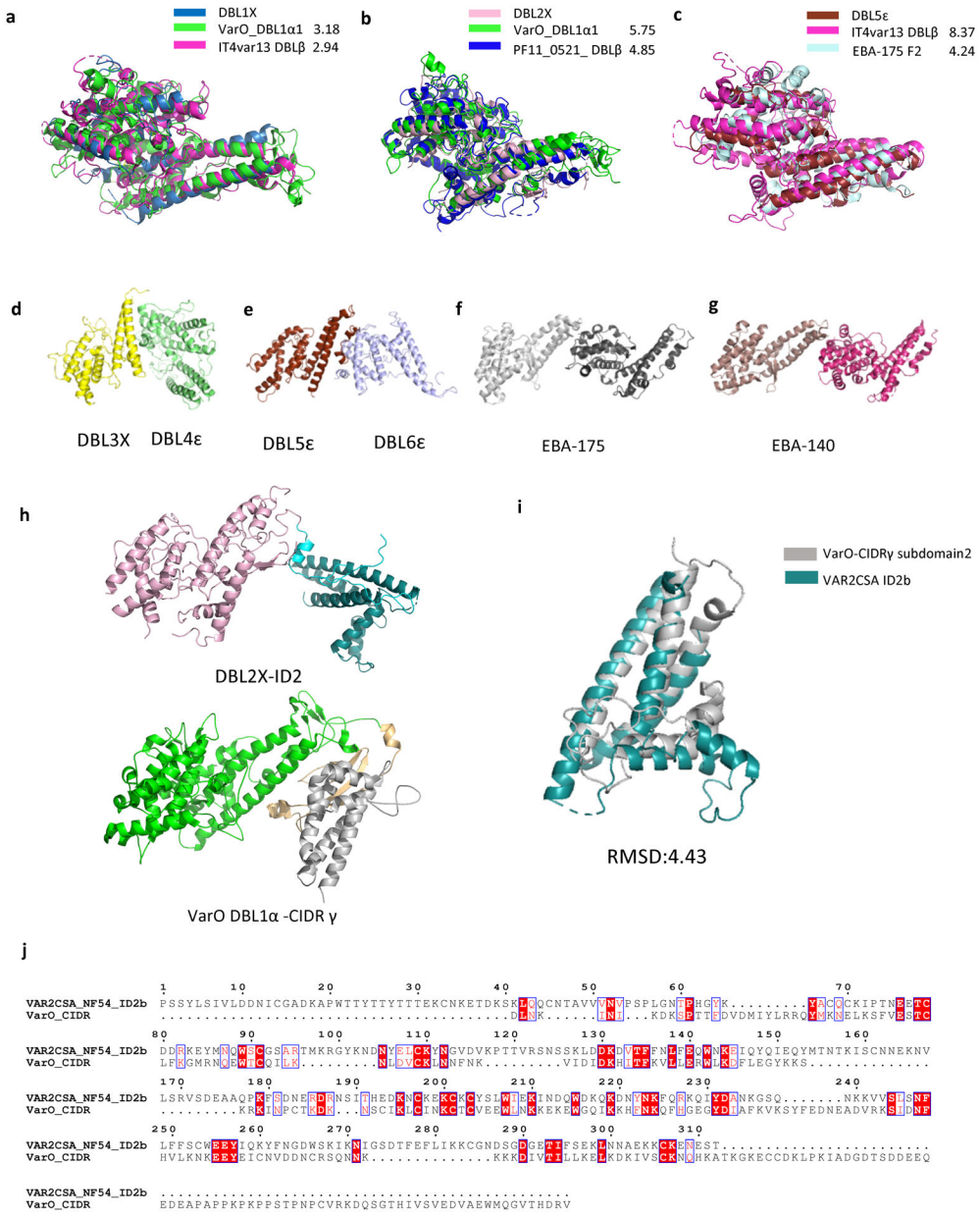
Extended Data Fig. 2. Data-processing pipeline for the cryo-EM movies of CSA-VAR2CSA NF54 complex.

a, Flow chart showing the image-processing pipeline for the cryo-EM data of VAR2CSA starting with 6,196 dose-fractionated movies collected on a 300-keV Titan Krios (FEI) equipped with a K2 Summit direct electron detector (Gatan). Data were processed in cryoSPARC. Full frame motion correction was done by cryoSPARC's own implementation. CTF estimation for each micrograph was calculated with Gctf. Particles were autopicked from each micrograph with the blob picker from cryoSPARC and then sorted by 2D classification for two rounds. The twelve highest-populated classes with clear features from the 2D classification are shown. The dataset contained 858,299 particles. A subset of particles was used to generate an ab initio map in cryoSPARC. Particles were classified into 5 classes using the low-pass-filtered (30 Å) ab initio map as a template. Class 1 was selected

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

with 157,702 particles to conduct NU-refinement and generated a 3.87 Å map. A mask covering DBL5ε and DBL6ε domains were then used to perform local refinement and generated a 4.88 Å map. Class 1 and Class 4 which have a clear core density were selected again with 299,571 particles to conduct NU-refinement and generated a 3.5 Å map, local refinement improved the resolution of the core to 3.36 Å. Scale bar: 10nm. Analysis was performed three times independently with similar results. b, Gold-standard FSC curves are shown. The dotted line represents the 0.143 FSC cut-off. c, Angular distribution calculated in cryoSPARC for particle projections of the full-length protein (right) and the core (left). Heat map shows number of particles for each viewing angle. d,e and f, Local resolution of the core (d), full length (e) and arm region (f) in two views. The representation of colors for different resolution are shown on the right. g, FSC calculated between the refined structures and the full map. h, Representative cryo-EM densities from the core machinery map. i, Representative cryo-EM densities from the arm with DBL5ε and DBL6ε model docked in.

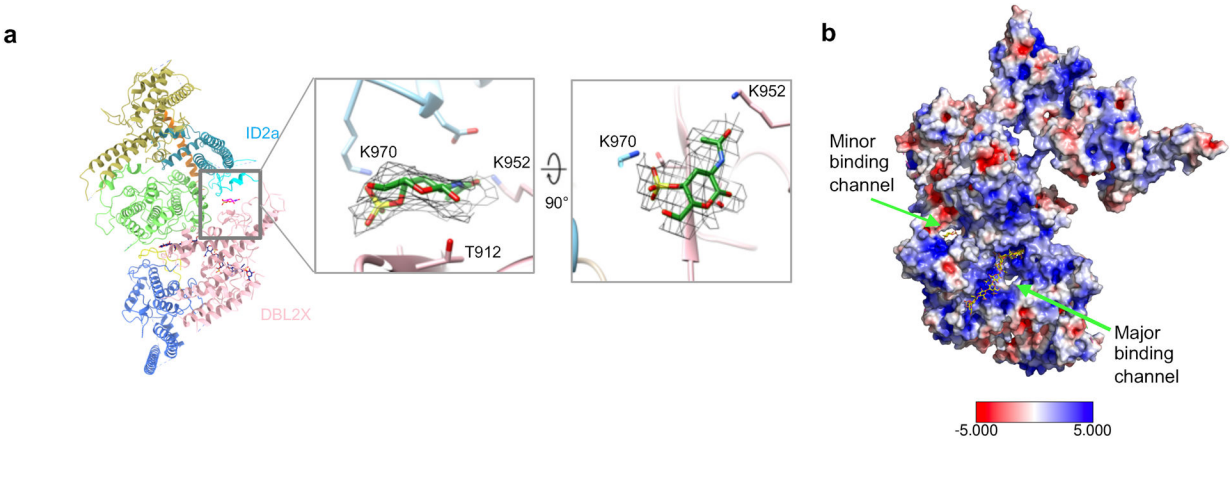
Author Manuscript Author Manuscript Author Manuscript Author Manuscript



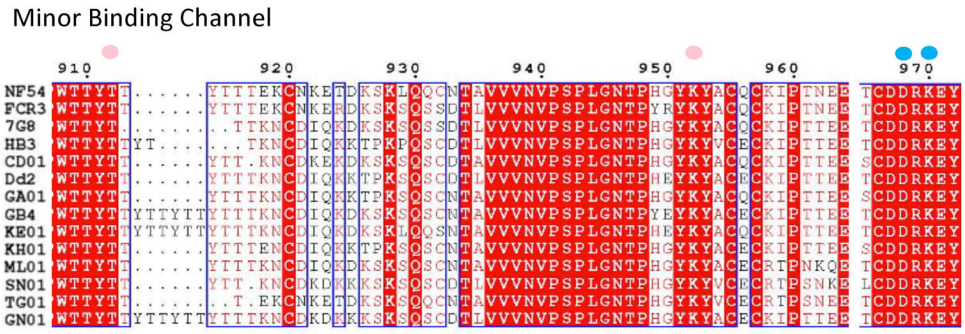
Extended Data Fig. 3. Structural conservation with the PfEMP1 family.

a, Structural alignment of DBL1X with VarO_DBL1α1 (PDB: 2YK0, RMSD:3.18) and IT4var13 DBLβ (PDB:6s8t, RMSD: 2.94). b, Structural alignment of DBL2X with varO_DBL1α1 (PDB: 2YKO, RMSD:5.75) and PF11_0521_ DBLβ (PDB: 5mza, RMSD:4.85). c, Structural alignment of DBL5ε with IT4var13 DBLβ (PDB: 6s8t, RMSD:8.37) and EBA-175 F2 domain (PDB: 1ZRO, RMSD:4.24). d, DBL3X-4ε and e, DBL 5ε−6ε, DBL domains are colored according to Fig. 1a. f, Crystal structure of EBA-175 (PDB: 1ZRO). The F1 and F2 domain are colored in light and dark grey respectively. g, Crystal structure of EBA-140 (PDB: 4JNO). The F1 and F2 domain are colored in brown and maroon respectively. h, Structural comparison of VAR2CSA DBL2X-ID and PfEMP1-VarO DBL1α-CIDR. Upper: atomic model of VAR2CSA DBL2X-ID2; Lower: Crystal

structure of varO_DBL1α1-CIDR γ. Green: DBL1α1, Gold: CIDR subdomain1, Grey: CIDR subdomain2. i, Structural alignment of VAR2CSA_ID2 with varO_CIDR γ subdomain2, RMSD=4.43. j, Sequence alignment of VAR2CSA ID2 and varO_CIDR γ.



c

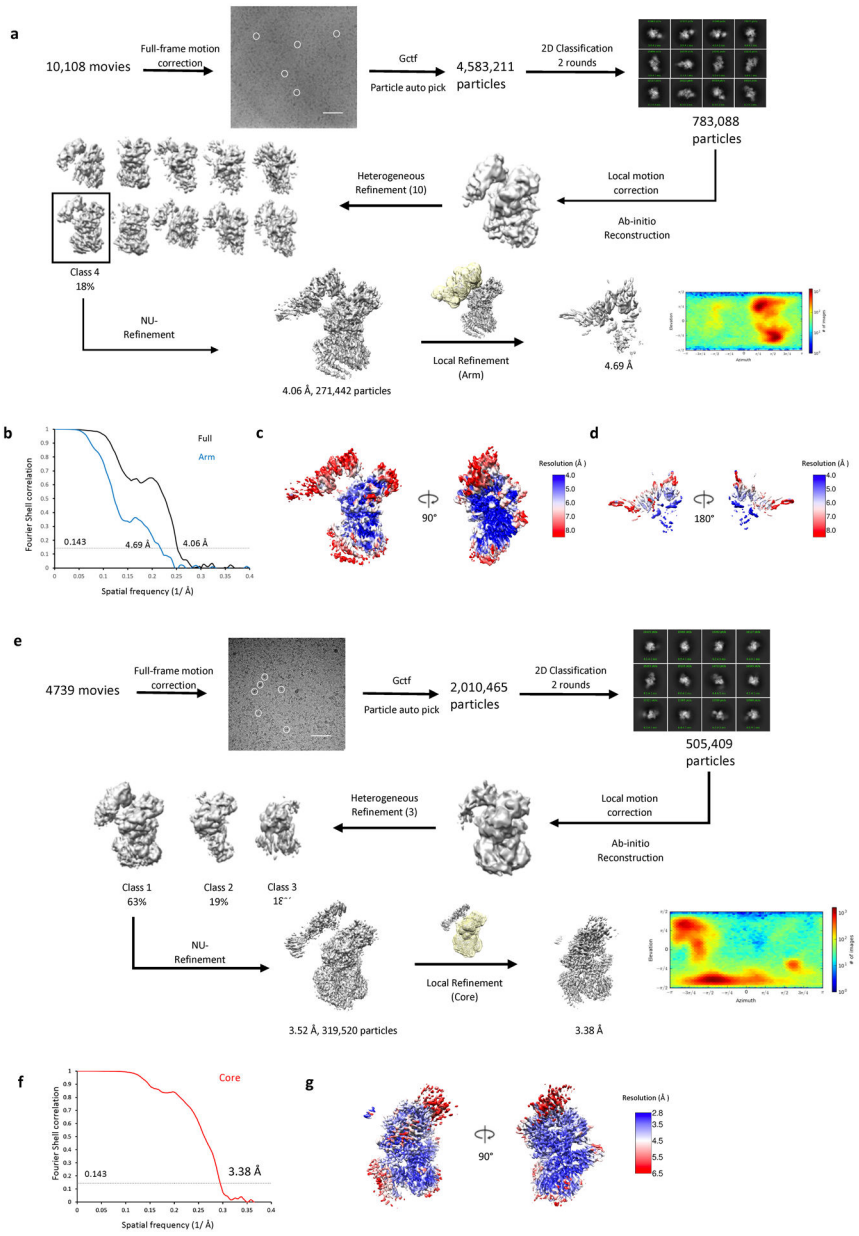


Extended Data Fig. 4. The recognition of CSA by VAR2CSA.

a, One ASG monosaccharide could be built in a weak density found in the minor binding channel sandwiched by DBL2X and ID2a. The density is shown in mesh. The ASG monosaccharide is colored in green. The residues that involve in forming hydrogen bonds with the ASG monosaccharide are illustrated. b, Electrostatic surface of the proteins showing both major binding channel and minor binding channel are positively charged. c, partial sequence alignment of the residues involved in the minor binding channel. The residues that interact with the monosaccharide from DBL2X and ID2a are highlighted on top by pink and blue spheres respectively.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



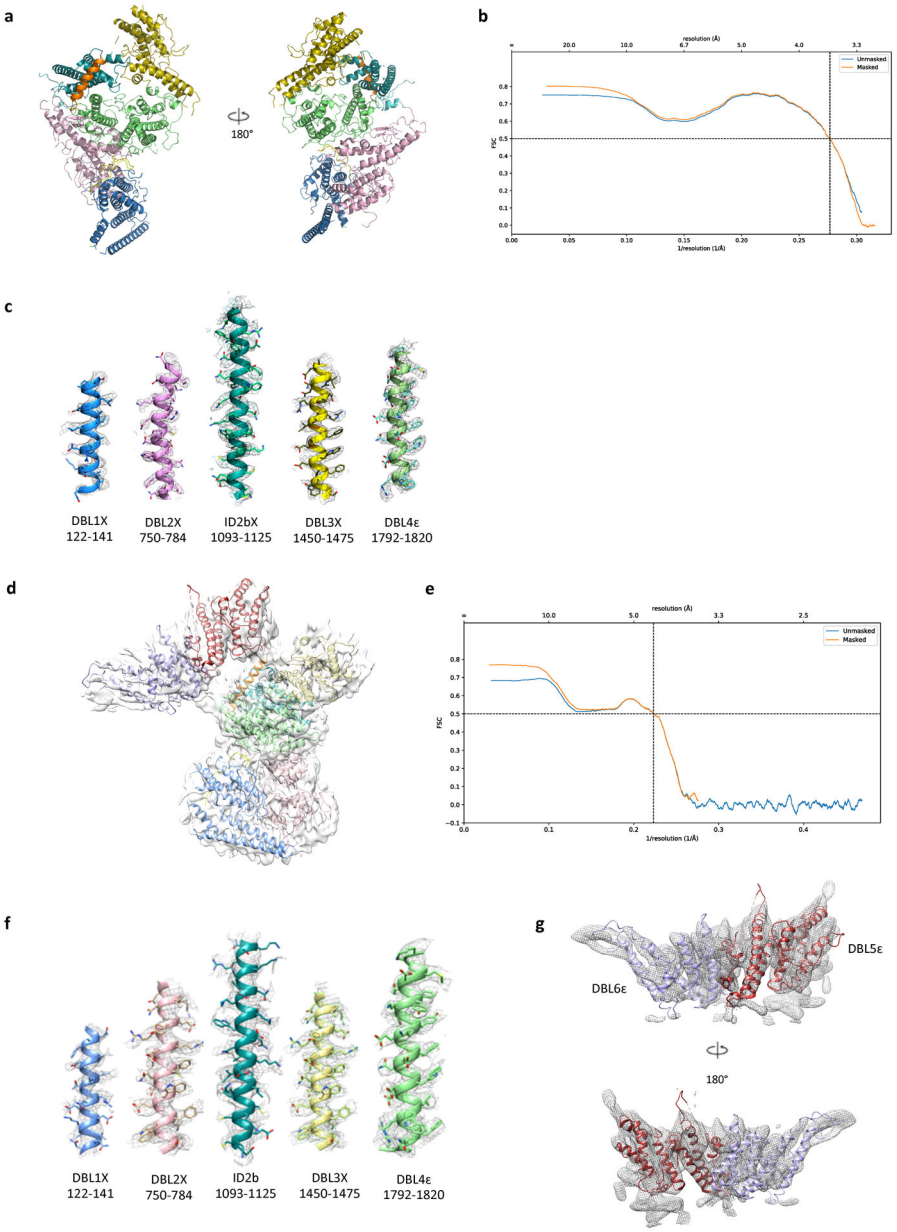
Extended Data Fig. 5. Data-processing pipeline for the cryo-EM movies of apo VAR2CSA FCR3 and VAR2CSA FCR3_crosslink.

a, Flow chart showing the image-processing pipeline for the cryo-EM data of VAR2CSA starting with 100,108 dose-fractionated movies collected on a 300-keV Titan Krios (FEI) equipped with a K2 Summit direct electron detector (Gatan). Processing was done within cryoSPARC. Full frame motion correction was done by cryoSPARC's own implementation. CTF estimation for each micrograph was calculated with Gctf. Particles were autopicked from each micrograph with the blob picker from cryoSPARC and then sorted by 2D classification for two rounds to exclude bad particles. The twelve highest-populated classes with clear features from the 2D classification are shown. The dataset contained 783,088 particles. A subset of particles was used to generate an ab initio map in cryoSPARC. Particles were classified into 10 classes using the low-pass-filtered (30 Å) ab initio map as a

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

template. Class 4 with a total of 271,442 particles was selected to conduct NU-refinement and generated a 4 Å map. A mask covering the arm region were then used to perform local refinement and generated a 4.69 Å map. The angular distribution calculated in cryoSPARC for particle projections are shown in heat map which shows number of particles for each viewing angle. Scale bar: 10nm. Analysis was performed three times independently with similar results. b, Gold-standard FSC curves. The dotted line represents the 0.143 FSC cut-off, which indicates a nominal resolution of 4 Å (black) and 4.69 Å (blue) for the full length protein and arm region respectively. c, Local resolution of the full length VAR2CSA map in two views. The representation of colors for different resolution are shown on the right. d, Local resolution of the arm map in two views. The representation of colors for different resolution are shown on the right. e, Flow chart showing the image-processing pipeline for the cryo-EM data of crosslinked VAR2CSA starting with 4,739 dose-fractionated movies collected on a 300-keV Titan Krios (FEI) equipped with a K2 Summit direct electron detector (Gatan). All processing was done within cryoSPARC. Full frame motion correction was done by cryoSPARC's own implementation and a sample. CTF estimation for each micrograph was calculated with Gctf. Particles were autopicked from each micrograph with the blob picker from cryoSPARC and then sorted by 2D classification for two rounds to exclude bad particles. The twelve highest-populated classes with clear features from the 2D classification are shown. The dataset contained 505,409 particles. A subset of particles was used to generate an ab initio map in cryoSPARC. Particles were classified into 3 classes using the low-pass-filtered (30 Å) ab initio map as a template. Class 3 was selected to conduct NU-refinement and generated a 3.52 Å map. A mask covering the core was then used to perform local refinement and generated a 3.38 Å map. The angular distribution calculated in cryoSPARC for particle projections are shown in heat map which shows number of particles for each viewing angle. Scale bar: 10nm. Analysis was performed three times independently with similar results. f, Gold-standard FSC curves. The dotted line represents the 0.143 FSC cut-off, which indicates a nominal resolution of 3.38 Å of the core. g, Local resolution of the crosslinked VAR2CSA FCR3 core map in two views. The representation of colors for different resolution are shown on the right.

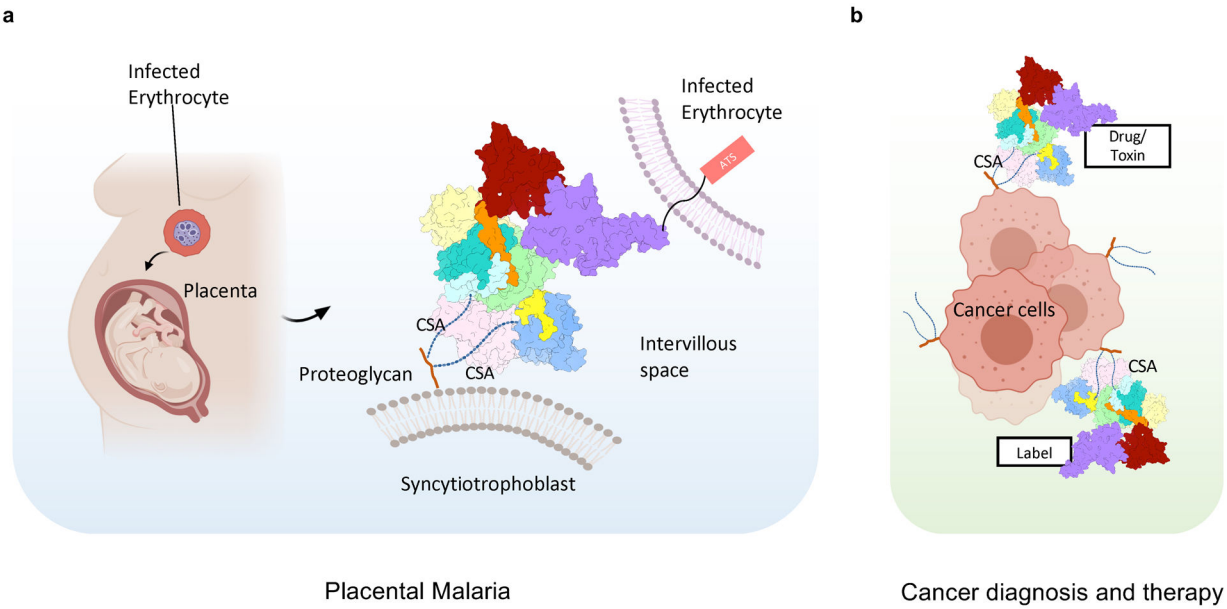
Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Extended Data Fig. 6. Model building and validation for VAR2CSA FCR3.

a, Atomic model of the core of the crosslinked VAR2CSA FCR3. b, FSC calculated between the refined structure and the full map. c, Representative cryo-EM densities from the core. d, Atomic model of full length VAR2CSA FCR3 docked in the 4.06 Å map. e, FSC calculated between the refined core structure and the full map. f, Representative cryo-EM densities from the core. g, Representative cryo-EM densities from the arm with DBL5ε and DBL6ε model docked in.

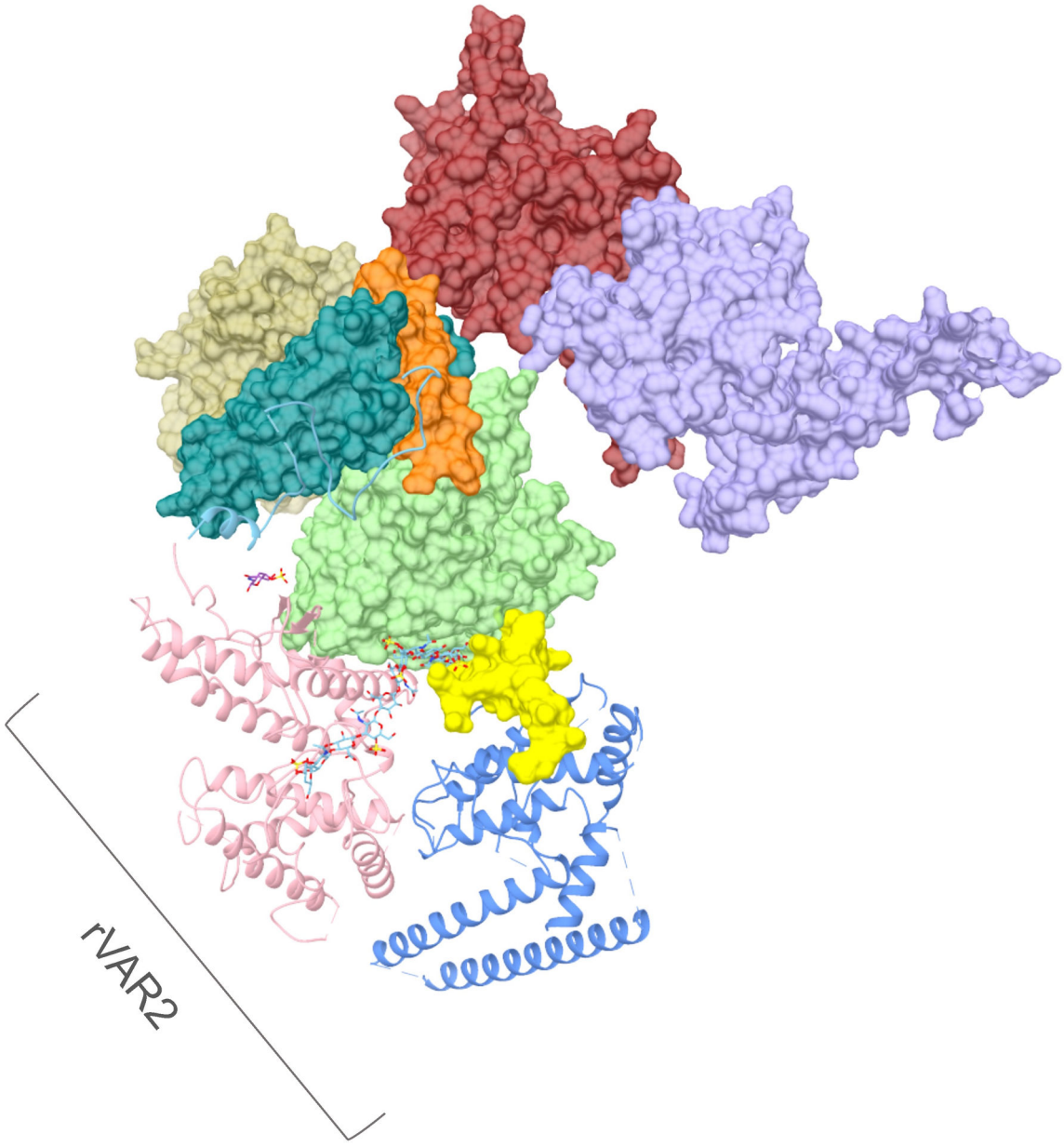
Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Extended Data Fig. 7. The model for VAR2CSA mediated placental malaria and cancer therapy.

a, The mechanism of placental sequestration of P. falciparum. In the placenta, the parasite express VAR2CSA on to the surface of the infected erythrocytes. VAR2CSA specifically binds to the CSA on the placental syncytiotrophoblast through a major and a potential minor CSA binding channel in its core with high affinity, leading to the sequestration of the parasite in the placenta and threaten the health of both the mother and their baby. b, Cancer cells of many cancer types harbor the same type of CSA on their surface as placenta. Conjugated VAR2CSA can be used to deliver drugs or labels specifically to tumor cells for therapeutics or diagnostics.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Extended Data Fig. 8. Analysis of rVAR2 for cancer therapy.

The structural model of rVAR2 are shown in ribbon. The remainder of the VAR2CSA protein is shown in surface. Each individual DBL and ID domain is colored according to Fig. 1a.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was funded by the Intramural Research Programs of NIAID and NICHD, National Institutes of Health. We thank NIH MICEF (Multi-Institute Cryo-EM Facility) and NCI/NICE (NIH IRP Cryo-EM Facility) for the data

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

collection. We sincerely thank J. Jiang for instrument access and expertise on sample preparation and data processing. We thank E. Fischer, V. Nair, C. Schwartz, B Hansen, A. Dearborn, and J. Marcotrigiano for their assistance in initial sample preparation and analysis. We thank N. Salinas, T. Dickey and W. Tang for their constructive suggestions on the manuscript. We thank J. Patrick Gorres for assistance in editing the manuscript. This study used the Office of Cyber Infrastructure and Computational Biology (OCICB) High Performance Computing (HPC) cluster at the National Institute of Allergy and Infectious Diseases (NIAID), Bethesda, MD.

References

1. Fried M & Duffy PE Malaria during Pregnancy. Cold Spring Harb Perspect Med 7, doi:10.1101/cshperspect.a025551 (2017).
2. Desai M et al. Epidemiology and burden of malaria in pregnancy. Lancet Infect Dis 7, 93–104, doi:10.1016/S1473-3099(07)70021-X (2007). [PubMed: 17251080]
3. Zakama AK, Ozarslan N & Gaw SL Placental Malaria. Curr Trop Med Rep, 1–10, doi:10.1007/s40475-020-00213-2 (2020).
4. Guyatt HL & Snow RW The epidemiology and burden of Plasmodium falciparum-related anemia among pregnant women in sub-Saharan Africa. Am J Trop Med Hyg 64, 36–44, doi:10.4269/ajtmh.2001.64.36 (2001). [PubMed: 11425176]
5. Steketee RW, Nahlen BL, Parise ME & Menendez C The burden of malaria in pregnancy in malaria-endemic areas. Am J Trop Med Hyg 64, 28–35, doi:10.4269/ajtmh.2001.64.28 (2001).
6. Duffy PE & Fried M Plasmodium falciparum adhesion in the placenta. Curr Opin Microbiol 6, 371–376, doi:10.1016/s1369-5274(03)00090-0 (2003). [PubMed: 12941407]
7. Ndam NT et al. Protective Antibodies against Placental Malaria and Poor Outcomes during Pregnancy, Benin. Emerg Infect Dis 21, 813–823, doi:10.3201/eid2105.141626 (2015). [PubMed: 25898123]
8. Pasternak ND & Dzikowski R PfEMP1: An antigen that plays a key role in the pathogenicity and immune evasion of the malaria parasite Plasmodium falciparum. Int J Biochem Cell B 41, 1463–1466, doi:10.1016/j.biocel.2008.12.012 (2009).
9. Salanti A et al. Evidence for the involvement of VAR2CSA in pregnancy-associated malaria. J Exp Med 200, 1197–1203, doi:10.1084/jem.20041579 (2004). [PubMed: 15520249]
10. Clausen TM et al. Structural and Functional Insight into How the Plasmodium falciparum VAR2CSA Protein Mediates Binding to Chondroitin Sulfate A in Placental Malaria. J Biol Chem 287, 23332–23345, doi:10.1074/jbc.M112.348839 (2012). [PubMed: 22570492]
11. Fried M & Duffy PE Designing a VAR2CSA-based vaccine to prevent placental malaria. Vaccine 33, 7483–7488, doi:10.1016/j.vaccine.2015.10.011 (2015). [PubMed: 26469717]
12. Trimnell AR et al. Global genetic diversity and evolution of var genes associated with placental and severe childhood malaria. Mol Biochem Parasitol 148, 169–180, doi:10.1016/j.molbiopara.2006.03.012 (2006). [PubMed: 16697476]
13. Bockhorst J et al. Structural polymorphism and diversifying selection on the pregnancy malaria vaccine candidate VAR2CSA. Mol Biochem Parasitol 155, 103–112, doi:10.1016/j.molbiopara.2007.06.007 (2007). [PubMed: 17669514]
14. Srivastava A et al. Full-length extracellular region of the var2CSA variant of PfEMP1 is required for specific, high-affinity binding to CSA. P Natl Acad Sci USA 107, 4884–4889, doi:10.1073/pnas.1000951107 (2010).
15. Bewley MC, Gautam L, Jagadeeshaprasad MG, Gowda DC & Flanagan JM Molecular architecture and domain arrangement of the placental malaria protein VAR2CSA suggests a model for carbohydrate binding. J Biol Chem, doi:10.1074/jbc.RA120.014676 (2020).
16. Dahlback M et al. The chondroitin sulfate A-binding site of the VAR2CSA protein involves multiple N-terminal domains. J Biol Chem 286, 15908–15917, doi:10.1074/jbc.M110.191510 (2011). [PubMed: 21398524]
17. Srivastava A et al. Var2CSA Minimal CSA Binding Region Is Located within the N-Terminal Region. Plos One 6, doi:ARTN e20270 10.1371/journal.pone.0020270 (2011). [PubMed: 21625526]
18. Resende M et al. Chondroitin sulphate A (CSA)-binding of single recombinant Duffy-binding-like domains is not restricted to Plasmodium falciparum Erythrocyte Membrane Protein 1 expressed by

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

CSA-binding parasites. Int J Parasitol 39, 1195–1204, doi:10.1016/j.ijpara.2009.02.022 (2009). [PubMed: 19324047]

19. Khunrae P, Philip JMD, Bull DR & Higgins MK Structural Comparison of Two CSPG-Binding DBL Domains from the VAR2CSA Protein Important in Malaria during Pregnancy. Journal of Molecular Biology 393, 202–213, doi:10.1016/j.jmb.2009.08.027 (2009). [PubMed: 19695262]
20. Babakhanyan A et al. The antibody response of pregnant Cameroonian women to VAR2CSA ID1-ID2a, a small recombinant protein containing the CSA-binding site. Plos One 9, e88173, doi:10.1371/journal.pone.0088173 (2014). [PubMed: 24505415]
21. Singh K et al. Structure of the DBL3x domain of pregnancy-associated malaria protein VAR2CSA complexed with chondroitin sulfate A. Nat Struct Mol Biol 15, 932–938, doi:10.1038/nsmb.1479 (2008). [PubMed: 19172746]
22. Higgins MK The structure of a chondroitin sulfate-binding domain important in placental malaria. J Biol Chem 283, 21842–21846, doi:10.1074/jbc.C800086200 (2008). [PubMed: 18550531]
23. Chene A et al. Down-selection of the VAR2CSA DBL1–2 expressed in E. coli as a lead antigen for placental malaria vaccine development. NPJ Vaccines 3, 28, doi:10.1038/s41541-018-0064-6 (2018). [PubMed: 30038803]
24. Chene A et al. Clinical development of placental malaria vaccines and immunoassays harmonization: a workshop report. Malar J 15, 476, doi:10.1186/s12936-016-1527-8 (2016). [PubMed: 27639691]
25. Sirima SB et al. PRIMVAC vaccine adjuvanted with Alhydrogel or GLA-SE to prevent placental malaria: a first-in-human, randomised, double-blind, placebo-controlled study. Lancet Infect Dis, doi:10.1016/S1473-3099(19)30739-X (2020).
26. Mordmuller B et al. First-in-human, Randomized, Double-blind Clinical Trial of Differentially Adjuvanted PAMVAC, A Vaccine Candidate to Prevent Pregnancy-associated Malaria. Clin Infect Dis 69, 1509–1516, doi:10.1093/cid/ciy1140 (2019). [PubMed: 30629148]
27. Salanti A et al. Targeting Human Cancer by a Glycosaminoglycan Binding Malaria Protein. Cancer Cell 28, 500–514, doi:10.1016/j.ccell.2015.09.003 (2015). [PubMed: 26461094]
28. Agerbaek MO, Bang-Christensen S & Salanti A Fighting Cancer Using an Oncofetal Glycosaminoglycan-Binding Protein from Malaria Parasites. Trends Parasitol 35, 178–181, doi:10.1016/j.pt.2018.11.004 (2019). [PubMed: 30551869]
29. Agerbaek MO et al. The VAR2CSA malaria protein efficiently retrieves circulating tumor cells in an EpCAM-independent manner. Nat Commun 9, doi:ARTN 3279 10.1038/s41467-018-05793-2 (2018). [PubMed: 30115931]
30. Nordor AV, Bellet D & Siwo GH Cancer-malaria: hidden connections. Open Biol 8, doi:10.1098/rsob.180127 (2018).
31. Tolia NH, Enemark EJ, Sim BK & Joshua-Tor L Structural basis for the EBA-175 erythrocyte invasion pathway of the malaria parasite Plasmodium falciparum. Cell 122, 183–193, doi:10.1016/j.cell.2005.05.033 (2005). [PubMed: 16051144]
32. Salinas ND, Tang WK & Tolia NH Blood-Stage Malaria Parasite Antigens: Structure, Function, and Vaccine Potential. J Mol Biol, doi:10.1016/j.jmb.2019.05.018 (2019).
33. Sim BK, Chitnis CE, Wasniowska K, Hadley TJ & Miller LH Receptor and ligand domains for invasion of erythrocytes by Plasmodium falciparum. Science 264, 1941–1944, doi:10.1126/science.8009226 (1994). [PubMed: 8009226]
34. Adams JH et al. A family of erythrocyte binding proteins of malaria parasites. Proc Natl Acad Sci U S A 89, 7085–7089, doi:10.1073/pnas.89.15.7085 (1992). [PubMed: 1496004]
35. Batchelor JD et al. Red blood cell invasion by Plasmodium vivax: structural basis for DBP engagement of DARC. PLoS Pathog 10, e1003869, doi:10.1371/journal.ppat.1003869 (2014). [PubMed: 24415938]
36. Lennartz F, Smith C, Craig AG & Higgins MK Structural insights into diverse modes of ICAM-1 binding by Plasmodium falciparum-infected erythrocytes. P Natl Acad Sci USA 116, 20124–20134, doi:10.1073/pnas.1911900116 (2019).
37. Lennartz F et al. Structure-Guided Identification of a Family of Dual Receptor-Binding PfEMP1 that Is Associated with Cerebral Malaria. Cell Host Microbe 21, 403–414, doi:10.1016/j.chom.2017.02.009 (2017). [PubMed: 28279348]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

38. Sander AF et al. Multiple var2csa-Type PfEMP1 Genes Located at Different Chromosomal Loci Occur in Many Plasmodium falciparum Isolates. Plos One 4, doi:ARTN e6667 10.1371/journal.pone.0006667 (2009). [PubMed: 19690615]

39. Holm L DALI and the persistence of protein shape. Protein Science 29, 128–140, doi:10.1002/pro.3749 (2020). [PubMed: 31606894]

40. Vigan-Womas I et al. Structural Basis for the ABO Blood-Group Dependence of Plasmodium falciparum Rosetting. Plos Pathogens 8, doi:ARTN e1002781 10.1371/journal.ppat.1002781 (2012). [PubMed: 22807674]

41. Malpede BM, Lin DH & Tolia NH Molecular basis for sialic acid-dependent receptor recognition by the Plasmodium falciparum invasion protein erythrocyte-binding antigen-140/BAEBL. J Biol Chem 288, 12406–12415, doi:10.1074/jbc.M113.450643 (2013). [PubMed: 23508963]

42. Achur RN, Valiyaveettil M & Gowda DC The low sulfated chondroitin sulfate proteoglycans of human placenta have sulfate group-clustered domains that can efficiently bind Plasmodium falciparum-infected erythrocytes. J Biol Chem 278, 11705–11713, doi:10.1074/jbc.M211015200 (2003). [PubMed: 12517756]

43. Achur RN, Valiyaveettil M, Alkhalil A, Ockenhouse CF & Gowda DC Characterization of proteoglycans of human placenta and identification of unique chondroitin sulfate proteoglycans of the intervillous spaces that mediate the adherence of Plasmodium falciparum-infected erythrocytes to the placenta. J Biol Chem 275, 40344–40356, doi:10.1074/jbc.M006398200 (2000). [PubMed: 11005814]

44. Agbor-Enoh ST et al. Chondroitin sulfate proteoglycan expression and binding of Plasmodium falciparum-infected erythrocytes in the human placenta during pregnancy. Infection and Immunity 71, 2455–2461, doi:10.1128/Iai.71.5.2455-2461.2003 (2003). [PubMed: 12704116]

45. Alkhalil A, Achur RN, Valiyaveettil M, Ockenhouse CF & Gowda DC Structural requirements for the adherence of Plasmodium falciparum-infected erythrocytes to chondroitin sulfate proteoglycans of human placenta. J Biol Chem 275, 40357–40364, doi:10.1074/jbc.M006399200 (2000). [PubMed: 11005815]

46. Gangnard S et al. Structure of the DBL3X-DBL4 epsilon region of the VAR2CSA placental malaria vaccine candidate: insight into DBL domain interactions. Sci Rep-Uk 5, doi:ARTN 14868 10.1038/srep14868 (2015).

47. Gangnard S et al. Structural and Immunological Correlations between the Variable Blocks of the VAR2CSA Domain DBL6 epsilon from Two Plasmodium falciparum Parasite Lines. Journal of Molecular Biology 425, 1697–1711, doi:10.1016/j.jmb.2013.02.014 (2013). [PubMed: 23429057]

48. Avril M et al. Evidence for globally shared, cross-reacting polymorphic epitopes in the pregnancy-associated malaria vaccine candidate VAR2CSA. Infect Immun 76, 1791–1800, doi:10.1128/IAI.01470-07 (2008). [PubMed: 18250177]

49. Glaser F et al. ConSurf: Identification of Functional Regions in Proteins by Surface-Mapping of Phylogenetic Information. Bioinformatics 19, 163–164, doi:10.1093/bioinformatics/19.1.163 (2003). [PubMed: 12499312]

50. Ditlev SB et al. Identification and Characterization of B-Cell Epitopes in the DBL4 epsilon Domain of VAR2CSA. Plos One 7, doi:ARTN e43663 10.1371/journal.pone.0043663 (2012). [PubMed: 22970138]

51. Doritchamou JYA et al. VAR2CSA Domain-Specific Analysis of Naturally Acquired Functional Antibodies to Plasmodium falciparum Placental Malaria. Journal of Infectious Diseases 214, 577–586, doi:10.1093/infdis/jiw197 (2016).

52. Barfod L et al. Human pregnancy-associated malaria-specific B cells target polymorphic, conformational epitopes in VAR2CSA. Mol Microbiol 63, 335–347, doi:10.1111/j.1365-2958.2006.05503.x (2007). [PubMed: 17176260]

53. Andersen P et al. Structural insight into Epitopes in the pregnancy-associated malaria protein VAR2CSA. Plos Pathogens 4, doi:ARTN e42 10.1371/journal.ppat.0040042 (2008). [PubMed: 18282103]

54. Mitran CJ et al. Antibodies to Cryptic Epitopes in Distant Homologues Underpin a Mechanism of Heterologous Immunity between Plasmodium vivax PvDBP and Plasmodium falciparum

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

VAR2CSA. Mbio 10, doi:ARTN e02343–19 10.1128/mBio.02343-19 (2019). [PubMed: 31594821]

55. Peters PJ, Thigpen MC, Parise ME & Newman RD Safety and toxicity of sulfadoxine/pyrimethamine: implications for malaria prevention in pregnancy using intermittent preventive treatment. Drug Saf 30, 481–501, doi:10.2165/00002018-200730060-00003 (2007). [PubMed: 17536875]
56. Patel JC et al. Increased risk of low birth weight in women with placental malaria associated with P. falciparum VAR2CSA clade. Sci Rep 7, 7768, doi:10.1038/s41598-017-04737-y (2017). [PubMed: 28801627]
57. Doritchamou JYA et al. Placental malaria vaccine candidate antigen VAR2CSA displays atypical domain architecture in some Plasmodium falciparum strains. Commun Biol 2, doi:ARTN 457 10.1038/s42003-019-0704-z (2019). [PubMed: 31840102]
58. Helms G, Dasanna AK, Schwarz US & Lanzer M Modeling cytoadhesion of Plasmodium falciparum-infected erythrocytes and leukocytes-common principles and distinctive features. Febs Lett 590, 1955–1971, doi:10.1002/1873-3468.12142 (2016). [PubMed: 26992823]
59. Cutts EE et al. Structural analysis of P. falciparum KAHRP and PfEMP1 complexes with host erythrocyte spectrin suggests a model for cytoadherent knob protrusions. Plos Pathogens 13, doi:ARTN e1006552 10.1371/journal.ppat.1006552 (2017). [PubMed: 28806784]
60. Dorin-Semblat D et al. Phosphorylation of the VAR2CSA extracellular region is associated with enhanced adhesive properties to the placental receptor CSA. Plos Biol 17, doi:ARTN e3000308 10.1371/journal.pbio.3000308 (2019). [PubMed: 31181082]
61. Doritchamou J et al. Differential adhesion-inhibitory patterns of antibodies raised against two major variants of the NTS-DBL2X region of VAR2CSA. Vaccine 31, 4516–4522, doi:10.1016/j.vaccine.2013.07.072 (2013). [PubMed: 23933341]
62. Akhouri RR, Goel S, Furusho H, Skoglund U & Wahlgren M Architecture of Human IgM in Complex with P. falciparum Erythrocyte Membrane Protein 1. Cell Rep 14, 723–736, doi:10.1016/j.celrep.2015.12.067 (2016). [PubMed: 26776517]
63. Brown A et al. Molecular Architecture of a Complex between an Adhesion Protein from the Malaria Parasite and Intracellular Adhesion Molecule 1. J Biol Chem 288, 5992–6003, doi:10.1074/jbc.M112.416347 (2013). [PubMed: 23297413]
64. Stevenson L et al. Investigating the function of F-c-specific binding of IgM to Plasmodium falciparum erythrocyte membrane protein 1 mediating erythrocyte rosetting. Cell Microbiol 17, 819–831, doi:10.1111/cmi.12403 (2015). [PubMed: 25482886]
65. Wahlgren M, Goel S & Akhouri RR Variant surface antigens of Plasmodium falciparum and their roles in severe malaria. Nat Rev Microbiol 15, 479–491, doi:10.1038/nrmicro.2017.47 (2017). [PubMed: 28603279]
66. Shukla AK et al. Visualization of arrestin recruitment by a G-protein-coupled receptor. Nature 512, 218–222, doi:10.1038/nature13430 (2014). [PubMed: 25043026]
67. Ohi M, Li Y, Cheng Y & Walz T Negative Staining and Image Classification - Powerful Tools in Modern Electron Microscopy. Biol Proced Online 6, 23–34, doi:10.1251/bpo70 (2004). [PubMed: 15103397]
68. Mastronarde DN Automated electron microscope tomography using robust prediction of specimen movements. Journal of Structural Biology 152, 36–51, doi:10.1016/j.jsb.2005.07.007 (2005). [PubMed: 16182563]
69. Suloway C et al. Automated molecular microscopy: The new Leginon system. Journal of Structural Biology 151, 41–60, doi:10.1016/j.jsb.2005.03.010 (2005). [PubMed: 15890530]
70. Punjani A, Rubinstein JL, Fleet DJ & Brubaker MA cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods 14, 290–296, doi:10.1038/nmeth.4169 (2017). [PubMed: 28165473]
71. Zhang K Gctf: Real-time CTF determination and correction. J Struct Biol 193, 1–12, doi:10.1016/j.jsb.2015.11.003 (2016). [PubMed: 26592709]
72. Afonine PV et al. New tools for the analysis and validation of cryo-EM maps and atomic models. Acta Crystallogr D Struct Biol 74, 814–840, doi:10.1107/S2059798318009324 (2018). [PubMed: 30198894]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

73. Kelley LA, Mezulis S, Yates CM, Wass MN & Sternberg MJE The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10, 845–858, doi:10.1038/nprot.2015.053 (2015). [PubMed: 25950237]

74. Emsley P & Cowtan K Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60, 2126–2132, doi:10.1107/S0907444904019158 (2004). [PubMed: 15572765]

75. Adams PD et al. PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr D Biol Crystallogr 58, 1948–1954, doi:10.1107/s0907444902016657 (2002). [PubMed: 12393927]

76. Pettersen EF et al. UCSF chimera - A visualization system for exploratory research and analysis. J Comput Chem 25, 1605–1612, doi:10.1002/jcc.20084 (2004). [PubMed: 15264254]

77. Whalen DM, Malinauskas T, Gilbert RJC & Siebold C Structural insights into proteoglycan-shaped Hedgehog signaling. P Natl Acad Sci USA 110, 16420–16425, doi:10.1073/pnas.1310097110 (2013).

78. Goddard TD et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci 27, 14–25, doi:10.1002/pro.3235 (2018). [PubMed: 28710774]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

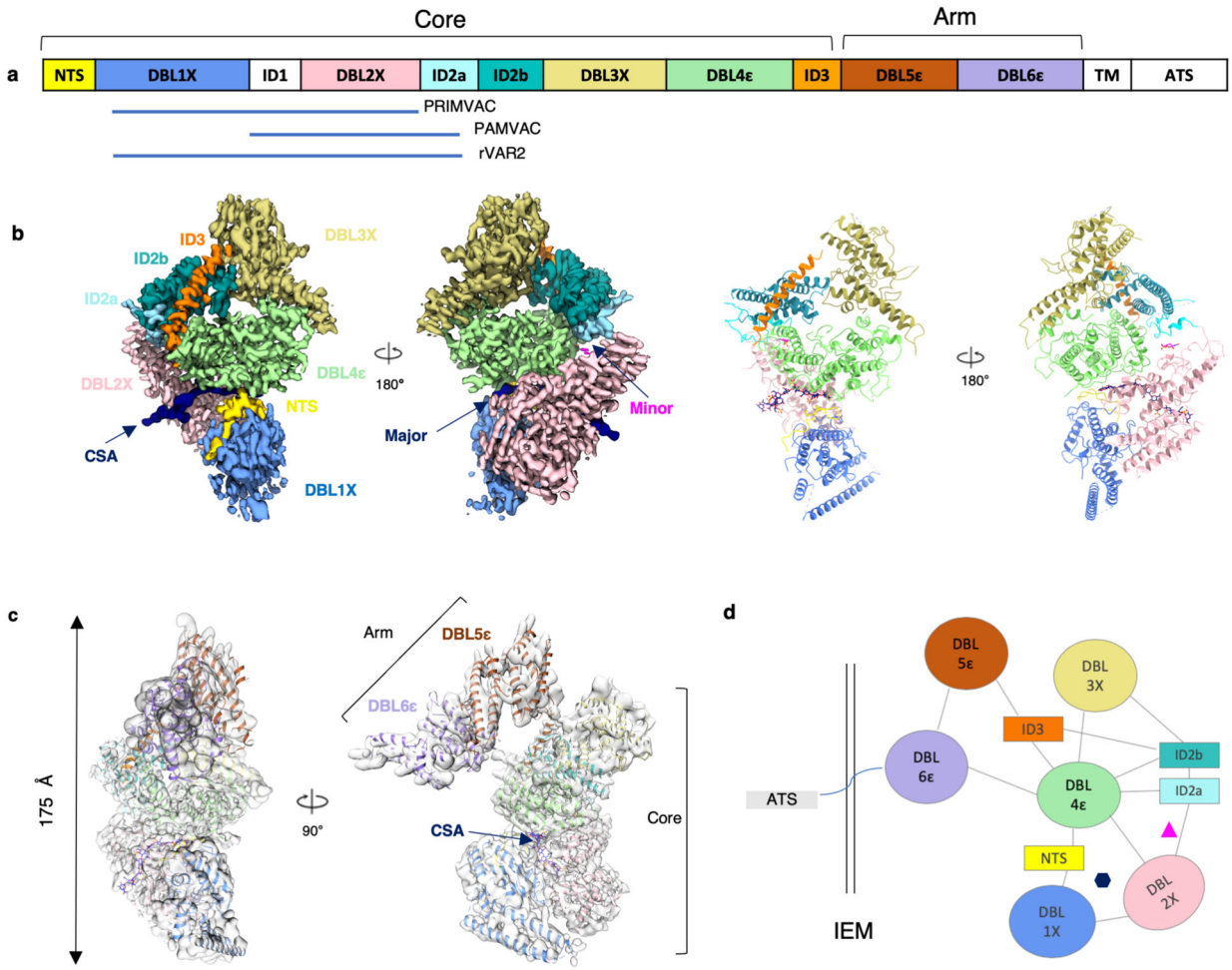


Fig. 1. Overall structure of the CSA-VAR2CSA NF54 complex.

a, Schematic of VAR2CSA NF54 primary structure colored by domain. Domains that were excluded from the ectodomain expression construct or could not be visualized in the final map are colored white. TM: transmembrane domain. ATS: Acidic terminal sequence. The alignments of PRIMVAC, PAMVAC and rVAR2 polypeptide are indicated below. **b,** Left: two views of the Cryo-EM density for the 3.36 Å core structure. Right: the same two views of the atomic model corresponding to the map. Each domain is colored as in **a**. The CSA major and minor binding channel are highlighted by arrows. The CSA polymer in the major binding channel is colored in dark blue and the CSA monosaccharide in the minor binding channel is colored in magenta. **c,** Two views of the Cryo-EM density for the entire complex with the model docked inside. The full-length density is the combination of the core and arm after local refinement. **d,** Schematic drawing of the CSA-VAR2CSA NF54 complex. Each line indicates interactions between the connecting domains. The major binding channel and minor binding channel are highlighted by the dark blue hexagon and magenta triangle.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

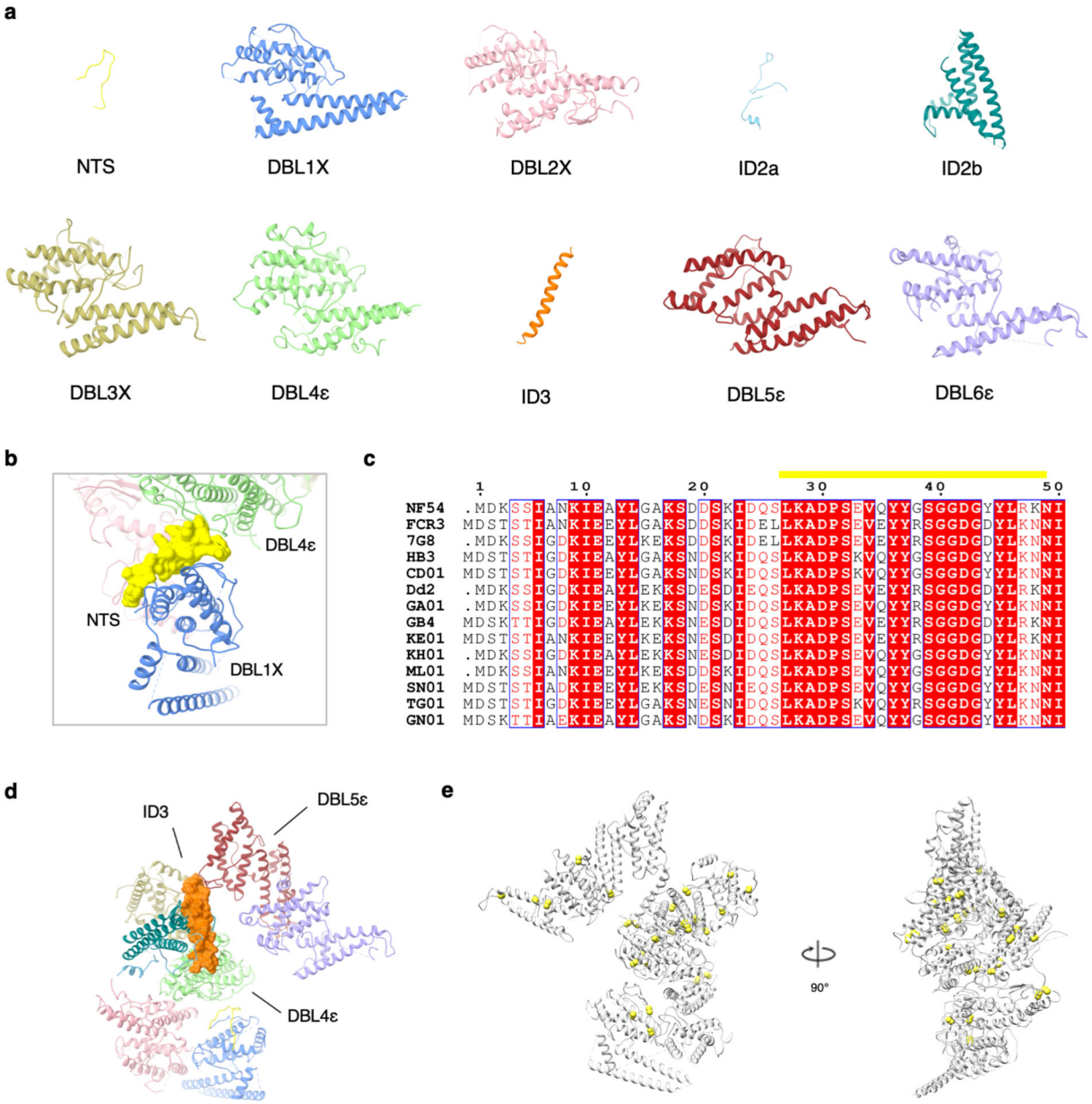


Fig.2. Domain composition of VAR2CSA.

a, The models of the NTS, six DBL domains and 2 ID regions are shown according to the order of the protein sequence. Each domain is colored according to Fig. 1a. **b,** NTS unites DBL1X and DBL4ε. NTS is shown in surface while DBL1X and DBL4ε are shown in ribbon. All the domains are colored according to Fig. 1a. **c,** Sequence alignment of NTS among different VAR2CSA variants. The range of the final model of the NTS is highlighted by the yellow bar above the sequences. **d,** ID3 is an α helix that connects DBL4ε and DBL5ε. ID3 is shown in surface while the rest of the molecule is shown in ribbon. All the domains are colored according to Fig. 1a. **e,** The models of VAR2CSA NF54 are colored in grey shown in two different views. The disulfides bonds are shown as yellow spheres.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

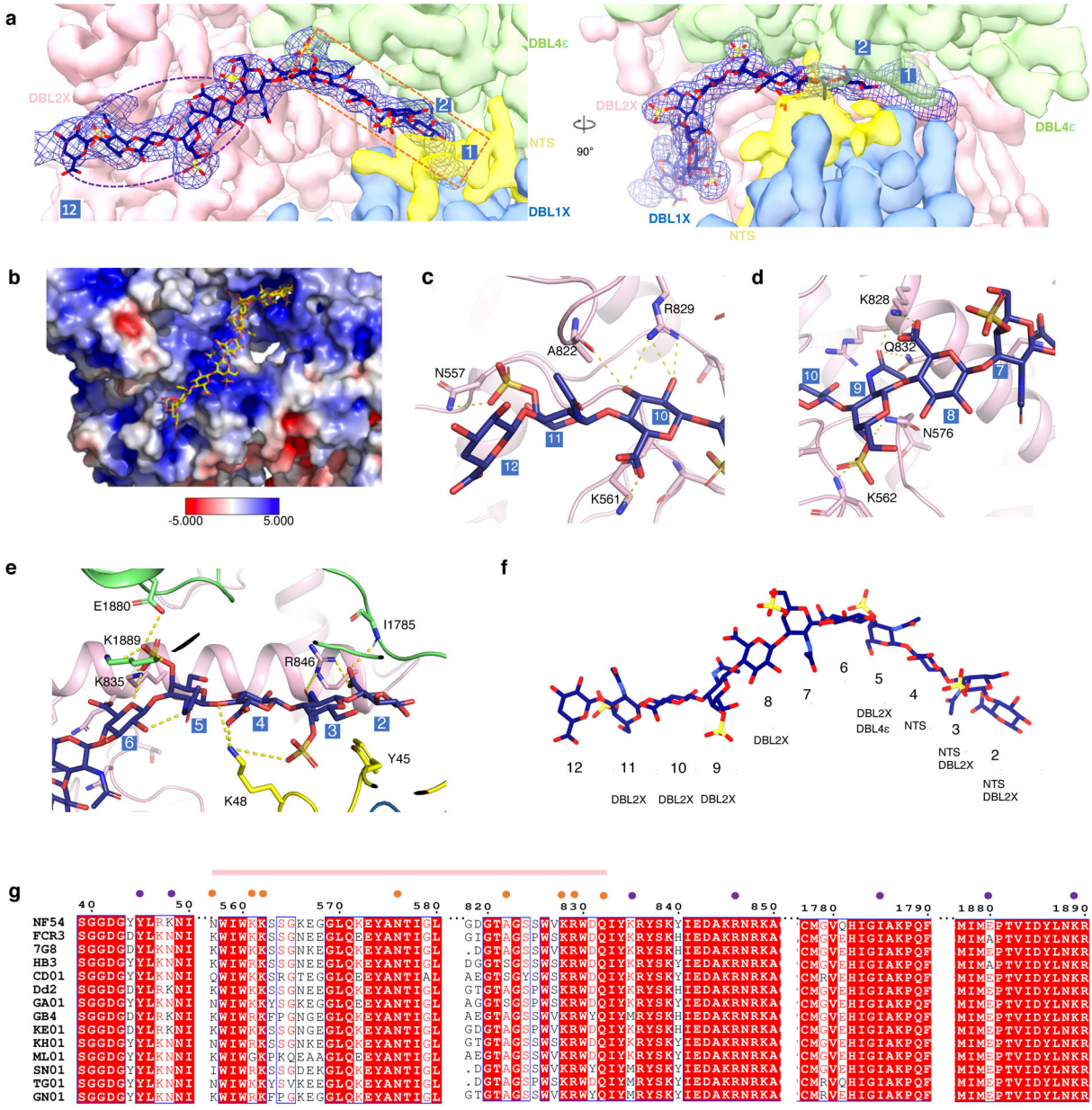


Fig. 3. CSA-binding sites within the major binding channel.

a, Two views of the structure showing a dodecamer of CSA is bound in the major binding channel. The Cryo-EM map density of NTS, DBL1X, DBL2X and DBL4ε are shown in solid with transparency. The Cryo-EM density of CSA is shown in mesh overlaid on the CSA model in stick. The left and right monosaccharide are BDP-12 and BDP-2 indicated by the numbers 12 and 2, respectively. Density for the first monosaccharide of the chain is also observed and labeled with 1. Binding site 1 and 2 are highlighted by purple oval and orange rectangle respectively. **b**, Electrostatic surface of VAR2CSA showing the positive charged binding channel of CSA. **c**,**d** Detailed interactions between CSA and binding site 1. Each monosaccharide is numbered. The protein sequence number and side chains of the residues involved in CSA recognition are shown. **e,** Detailed interactions between BDP-2 to ASG-5

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and binding site 2. Each monosaccharide is numbered. The protein sequence number and side chains of the residues involved in the CSA recognition are shown. **f,** The CSA molecule in the major binding channel is positioned as in Fig. 2A with numbering of each of the monosaccharide. The domains that each monosaccharide interact are indicated below. **g,** partial sequence alignment of the residues involved in binding CSA in the major binding channel, the residues in major binding sites 1 and 2 are highlighted on top by orange and purple spheres respectively. The surface exposed binding site on DBL2 is highlighted by the pink line.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

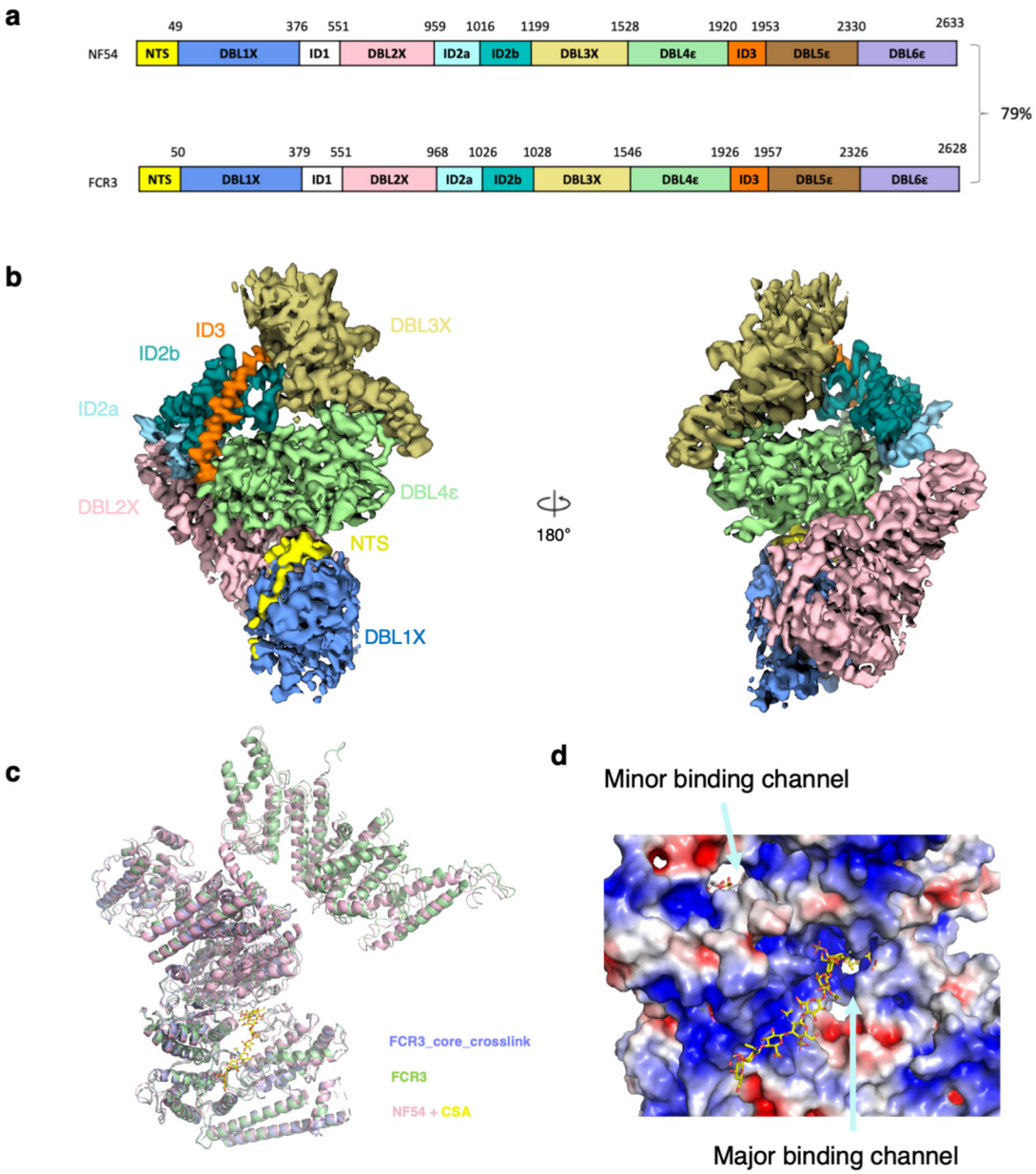


Fig. 4. VAR2CSA adopts preformed CSA binding channels.

a, Domain boundaries of VAR2CSA NF54 and VAR2CSA FCR3 ectodomains. The protein sequence identity between the two is labeled. **b,** Two views of the Cryo-EM density for the 3.38 Å core region of VAR2CSA FCR3. **c,** Structural alignment of apo VAR2CSA FCR3, cross-linked VAR2CSA FCR3 and the CSA-VAR2CSA NF54 complex. VAR2CSA NF54 is in purple, CSA in yellow, VAR2CSA FCR3 in yellow, and VAR2CSA FCR3 crosslinked in blue. **d,** The Electrostatic surface of VAR2CSA FCR3 is shown on the left with a zoom-in view of the CSA binding sites on the right. The major and minor binding channel are indicated by arrows.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

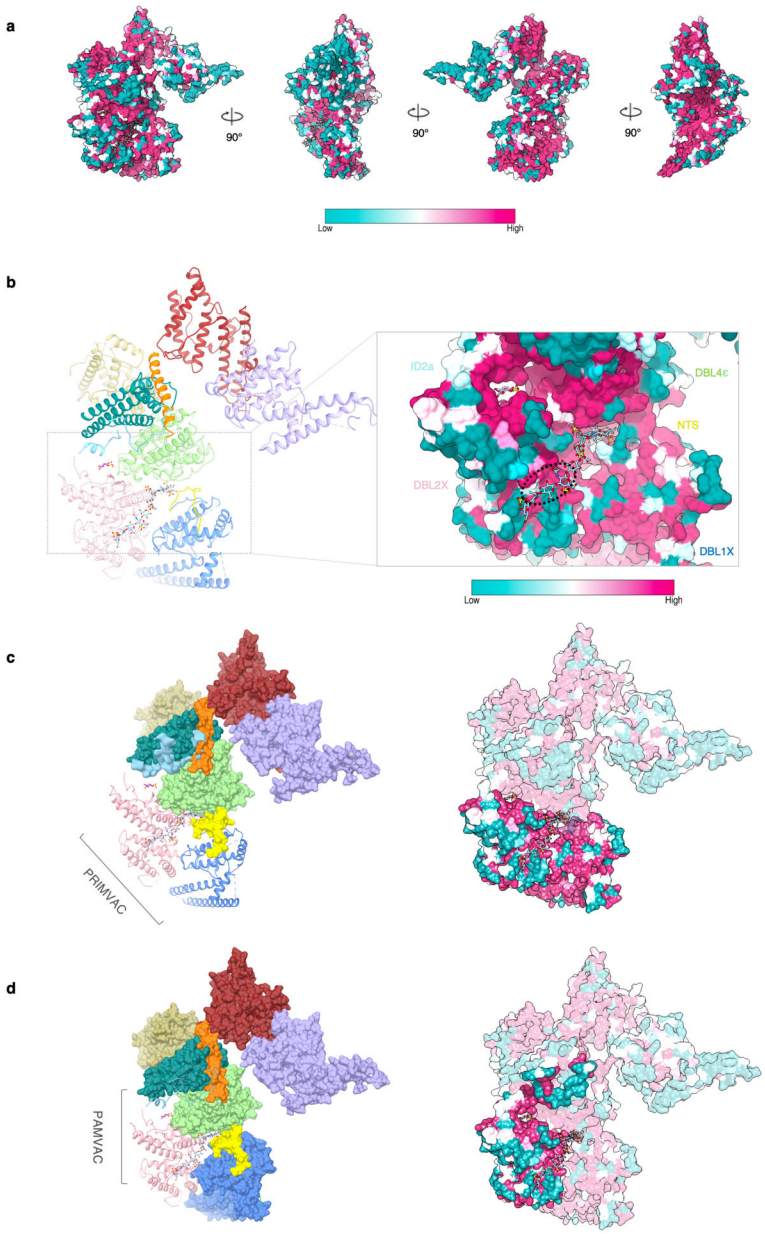


Fig. 5. Variability analysis of VAR2CSA.

a, Fourteen sequences of VAR2CSA that represent the diversity were analyzed using ConSurf[49]. Surface residues on a space filled model are shaded according to degree of conservation. The color key is shown below. Four different views are illustrated. **b,** Left: the atomic model of CSA-VAR2CSA NF54 complex. Right: space filling models of the CSA-VAR2CSA binding interface. Surface residues are shaded according to degree of conservation. The color key is shown below. The surface exposed major binding site 1 is highlighted by a black dotted circle. **c,** Left: the structural model of sequences comprising PRIMVAC and are shown in ribbon. The remainder of the VAR2CSA protein is shown in surface. Right: based on the variability analysis in **a**, PRIMVAC is shown in bold while the rest of the VAR2CSA molecule is shown in transparent. **d,** Left: the structural model of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

sequences comprising PAMVAC and are shown in ribbon. The remainder of the VAR2CSA protein is shown in surface. Right: based on the variability analysis in **a**, PAMVAC are shown in bold while the rest of the VAR2CSA molecule is shown in transparent.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

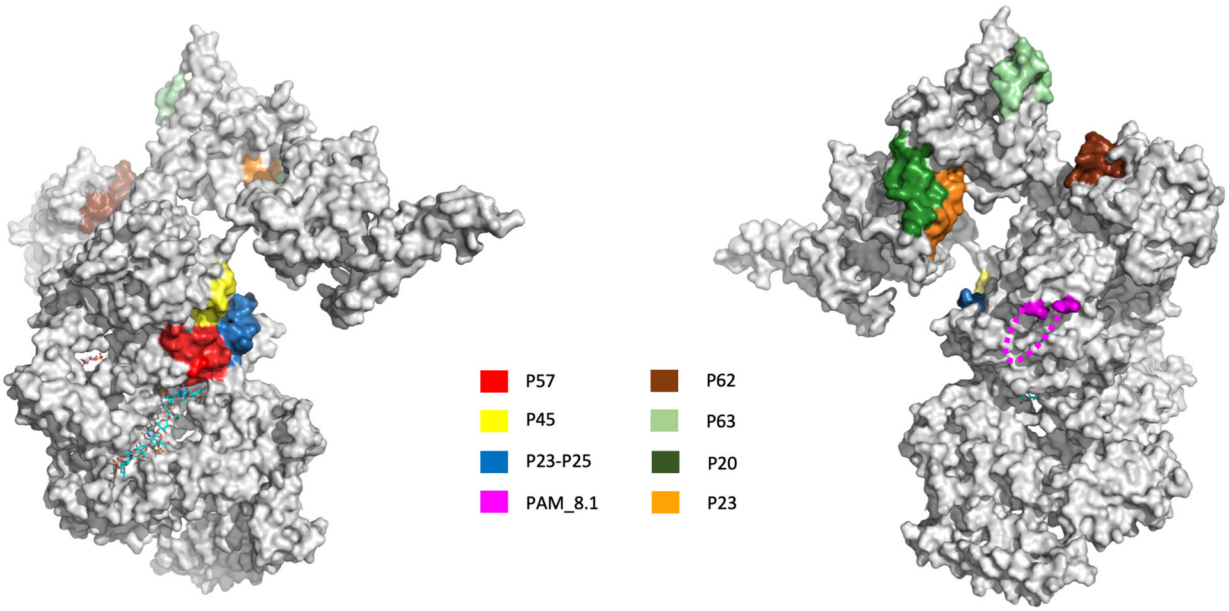


Fig. 6. Human antibodies epitopes mapped on VAR2CSA.

VAR2CSA structure is shown in surface. The characterized peptide epitopes are colored as illustrated: P57 (red), P54 (yellow) and P23-P25 (blue) are epitopes on DBL4ε. PAM8.1 epitope is a flexible loop on DBL3X that is missing the final structure and is colored by pink and illustrated by dash line. P62 on DBL3X is shown in brown and P63 on DBL5ε is shown in light green. The cryptic epitopes P20 and P23 on DBL5ε are shown in dark green and orange respectively.