INAUGURAL - DISSERTATION

zur
Erlangung der Doktorwürde
der
Naturwissenschaftlich-Mathematischen Gesamtfakultät
der
Ruprecht - Karls - Universität
Heidelberg

vorgelegt von

Master of Science Katharina Christina Imkeller

aus München

Tag der mündlichen Prüfung: .....

Molecular characterization of public anti-*Pf*CSP antibodies in human malaria

Gutachter:

Prof. Dr. Thomas Höfer

Prof. Dr. Hedda Wardemann

1 Abbreviations

7-AAD	7-aminoactinomycin D
ABTS	2,2'-azino-bis(3-ethylbenzthiazoline-6-sulphonic acid)
Ab	antibody
Ag	antigen
AID	activation induced deaminase
AIRR	adaptive immune receptor repertoire
AtM	atypical B cell memory population
AUC	area under curve
BCR	B cell receptor
BSA	bovine serum albumin
CD	cluster of differentiation
cDNA	complementary DNA
CDR	complementarity determining region
CHMI	controlled human malaria infection
CM	classical B cell memory population
CSP	circumsporozoite protein
DNA	deoxyribonucleic acid
dNTP	deoxynucleotid triphosphate
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EEF	exoerythrocytic form
ELISA	enzyme-linked immunosorbent assay
FACS	fluorescence-activated cell sorting
Fab	antigen binding fragment
Fig.	Figure
FITC	fluorescein isothiocyanate
FSC	forward scatter
FWR	framework region
HC-04 cells	human hepatocyte cells
HEK cells	human embryonic kidney cells
HIV	Human immunodeficiency virus
HRP	horseradish peroxidase
Ig	immunoglobulin
LaCHMI	Lambaréné controlled human malaria infection
LB	lysogeny broth
MN	mature naive (B cell)

NMR	nuclear magnetic resonance
Pb	*Plasmodium berghei*
PB	plasmablast
PBMC	peripheral blood mono-nuclear cell
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
Pf	*Plasmodium falciparum*
PFA	paraformaldehyde
Pb	*Plasmodium berghei*
Py	*Plasmodium yoelii*
RHP	random hexamer primer
RNA	ribonucleic acid
RT	reverse transcription
SHM	somatic hypermutation
SSC	side scatter
TuCHMI	Tübingen controlled human malaria infection

2 Table of contents

3 Zusammenfassung

Plasmodium falciparum ist ein einzelliger Parasit, der im Laufe seines komplexen Lebenszyklus *Anopheles* Mücken und Menschen befällt. Sporozoiten sind die Parasitenstadien, die von der Mücke in den Menschen gelangen. Sie können durch vom menschlichen Immunsystem generierte Antikörper neutralisiert werden. In dieser Studie wird die humorale Immunantwort, d.h. die Antikörperantwort, des Menschen auf das wichtigste Oberflächenantigen der Sporozoiten untersucht, das Circumsporozoit Protein (CSP).

Eine kürzlich veröffentliche Studie beschreibt hochaffine CSP-Antikörper, die im Laufe einer Reihe von kontrollierten humanen Malariainfektionen in europäischen Donoren entstehen, die dem Parasiten zuvor noch nie ausgesetzt waren. In der vorliegenden Arbeit werden zum einen die molekularen Eigenschaften ermittelt, die der Bindung dieser Antikörper zugrundeliegen. Außerdem wird untersucht, ob afrikanische Donoren aus einem endemischen Malariagebiet ähnliche Antikörper besitzen.

Im Repertoire der europäischen Donoren wurden mehrere Antikörpergruppen identifiziert, die sehr ähnliche Sequenzeigenschaften und Bindungseigenschaften besitzen, obwohl sie aus unterschiedlichen Spendern stammen. Es wird gezeigt, dass diese Antikörper nur an das Antigen binden können, wenn präzise Sequenzeigenschaften eingehalten werden. Dies reduziert vermutlich die Anzahl möglicher B Zell Vorläufer, die zur Entstehung der Antikörper führen können. Die wichtigste Erkenntnis dieser Studie ist die Beobachtung, dass Antikörper, die den repetitiven Bereich von CSP binden, mit sich selbst interagieren und auch Affinitätsreifung gegen die Antigenbindungsoberfläche des gegebüberliegenden Antikörpers aufweisen. Erstaunlicherweise konnten in den drei untersuchten afrikanischen Donoren keine Antikörper mit messbarer CSP-Reaktivität identifiziert werden. Die Arbeit liefert Anhaltspunkte dafür, dass dies zum einen an der Seltenheit der spezifischen B Zell Vorläufer liegt, aber auch am Bindungsmodus, der durch die repetitive Struktur des Antigens induziert wird.

4 Abstract

Plasmodium falciparum is a unicellular parasite that throughout its complex lifecycle infects *Anopheles* mosquitos and humans. The parasite stages injected into the human by the mosquito are called sporozoites and can be neutralized by antibodies generated by the human immune system. The present study investigates the humoral immune response, i.e. the antibody response, to the major antigen on the sporozoite surface, circumsporozoite protein (CSP).

A recently published study describes high-affine anti-CSP antibodies, which are generated upon repeated controlled human malaria infection of European donors, who were never exposed to the parasite before. The present work on the one hand describes the molecular characteristics that determine the binding of such antibodies. On the other hand, it investigates whether African donors from an endemic malaria region exhibited similar antibodies.

In the repertoire of the European donors, several groups of antibodies were identified that share highly similar amino acid sequences and binding behaviours. It is shown that these antibodies can only bind to the antigen if very specific sequence characteristics are conserved. This probably restricts the number of potential B cell precursors that can lead to the generation of such antibodies. The most important result of this study is the observation that antibodies binding the repetitive region of CSP directly interact with each other. The antibodies even show signs of anti-idiotope affinity maturation directed against the antigen binding site of neighbouring antibodies. Surprisingly in none of the three probed African donors antibodies with measurable anti-CSP reactivity could be found. The present work suggests that this is due to the low frequency of B cell precursors, but also due to the specific binding mode that is induced by the repetitive CSP structure.

5 Introduction

5.1 Circumsporozoite protein as a target of pre-erythrocytic immunity to *Plasmodium falciparum*

5.1.1 *Plasmodium falciparum* life cycle and infection of humans

Plasmodium falciparum (*Pf*) is a unicellular protozoan organism that parasitizes its primary mosquito and secondary human host. As all parasites from the Apicomplexan phylum, *Pf* undergoes major phenotypic changes throughout the different stages of its life cycle. The primary host of *Pf* is the female *Anopheles* mosquito, which takes up female and male gametocytes from an infected human host during a blood meal (Figure 1A). In the mosquito mid-gut the parasites mate, undergo sexual developmental stages and finally divide and differentiate into sporozoites, which invade the mosquito salivary gland. During a second blood meal these sporozoites are transmitted into the skin of another human host (Figure 1B). From the site of injection, the sporozoites migrate to the liver to invade hepatocytes, the cell type in which they differentiate into exoerythrocytic forms (EEFs). The EEFs grow, divide and finally differentiate into thousands of first generation merozoites. The latter are released into the blood stream and undergo multiple rounds of erythrocyte invasion, replication and cell lysis leading to an exponential growth in parasite numbers (Figure 1C). Some of the merozoites eventually differentiate into gametocytes, which can infect a mosquito and start another round of sexual development.

The blood stage replication of *Pf* in the human host, also called blood stage parasitemia, is the symptomatic part of the human malaria tropica disease and is accompanied by fever, vomiting, headaches, and tiredness. It is generally detected by blood smear, staining, and microscopical detection of blood stages inside the red blood cells. The disease is life threatening if untreated. The currently available chemotherapeutics against malaria are mostly effective against the blood stages of the parasite. Chloroquine for example interferes with the haemoglobin metabolism of merozoites in red blood cells. Resistance against chloroquine has been described in endemic areas. It is thus sometimes replaced by a combination therapy of lumefantrine and the artemisinin derivate artemether (Coartem), which is less susceptible to drug resistance.

5.1.2 Pre-erythrocytic immunity against *Plasmodium falciparum* sporozoites

The successful migration to and infection of hepatocytes by the injected sporozoites is a crucial step in human malaria and represents a bottleneck in the parasite life cycle. Only approximately 20 *Pf* sporozoites are injected into the human skin during a blood meal [11, 87]. It has been postulated that small numbers of parasites or even single surviving *Pf* sporozoites can be responsible for breakthrough of blood stage infection [121].

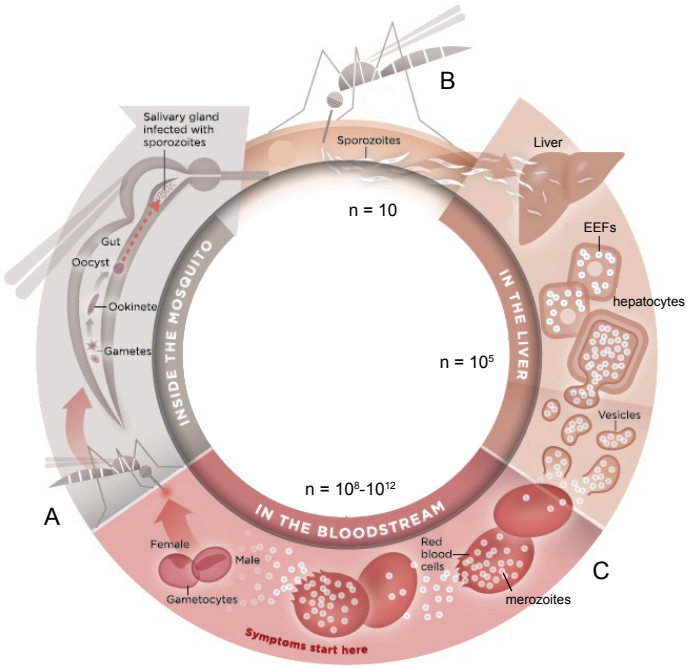


Figure 1. Life cycle of *Plasmodium falciparum*. A female *Anopheles* mosquito takes up gametocytes from an infected human host during a blood meal (A). Sporozoites develop in the mosquito and are injected into the human skin from where they migrate to the liver (B). Inside the hepatocytes the parasites differentiate to merozoites that are released to the blood stream and cause symptomatic blood stage parasitemia (C). Approximate parasite numbers at the different stages are indicated as n. Modified from MVI PATH [77].

Indeed the sporozoite biology is optimized for efficient migration and infection once the parasite is deposited in the human host. Investigations on the rodent-infecting *Plasmodium berghei* (*Pb*) and *Plasmodium yoelii* (*Py*) parasites have shown that after injection into the skin, sporozoites utilize gliding locomotion to propagate there and eventually reach blood or lymphatic vessels [5, 108]. Parasites ingressing lymphatic vessels are drained to the lymph nodes and mostly degraded [5]. The parasites reaching the blood stream are carried through the body and access the hepatic tissue via Kupffer cells [83]. Inside the liver they first traverse several layers of cells [62, 65] before they invade a hepatocyte and form a parasitophorous vacuole in which they differentiate [129]. The complete journey between injection and hepatocyte invasion takes only a few hours [32].

Naturally acquired humoral immune responses to the pre-erythrocytic parasite stages are very rare [46, 104], and seroepidemiological studies failed to prove correlation between pre-erythrocytic immunity and protection [41, 117]. In endemic areas protection rather correlates with immune responses to blood stage parasites [31, 40, 74]. It has even been described that semi-immune individuals carry small parasite numbers in the blood, a phenomenon called asymptomatic parasitemia. Immunity against blood stage parasites is non-sterile and does not disrupt the parasite spreading, because semi-immune individuals can serve as parasite reservoirs.

5.1.3 Controlled human malaria infection as a vaccination approach

With the aim of inducing sterile malaria immunity, research has focused on the investigation of controlled infection with live sporozoites as a potential vaccination approach to generate pre-erythrocytic immunity.

Controlled malaria infection is a method that has been used already in the early 20th century, first as a treatment against neurosyphilis, later also as an experimental setup to test the effectiveness of antimalarial drugs. From the 1960s on investigators started circumventing the symptomatic blood stage parasitemia by sporozoite attenuation or drug administration and thereby focused the immune response on the pre-erythrocytic stages. In early studies mosquitos carrying *P. berghei* sporozoites were x-ray-irradiated and then used to infect mice. Development of protective immunity was observed in these experiments [72]. In subsequent studies with humans, mosquitos carrying *Pf* sporozoites were irradiated [19, 42] and then used to infect study participants. Later also non-attenuated living sporozoites were used to infect humans by intra-dermal administration. The blood stage development of *Pf* was impeded by chloroquine chemoprophylaxis [9]. It was confirmed that the protection after controlled malaria infection is mainly mediated by pre-erythrocytic immunity [13]. Antibodies and memory B cells against pre-erythrocytic targets are induced in mice [43, 81, 130] and humans [10, 70]. It is nevertheless difficult to correlate them with protection against parasite infection.

Very recent studies utilizing intravenously injected live *Pf* sporozoites as immunogen in a vaccination approach (Sanaria PfSPZ-CVac) have demonstrated complete protection after a three repeated controlled infections [64]. During the course of the vaccination patients are under chloroquine chemoprophylaxis and the parasite thus cannot initiate the symptomatic infection of red blood cells. In addition the number of injected sporozoites is in general by orders of magnitude higher than the one injected during a mosquito blood meal, providing the immune system with large amounts of antigen. The protein based pre-erythrocytic vaccine RTS,S/AS01 (Mosquirix) is less effective and only reaches 25-50% vaccination efficacy in infants and young children [89, 90].

5.1.4 Circumsporozoite protein is one of the major pre-erythrocytic antigens

Circumsporozoite protein (CSP) is the major antigen on the mature sporozoite surface. The protein was first described and isolated in *Pb* as Pb44 and proven to be important for parasite infectivity [4, 81, 128]. Indeed, CSP is important in the context of ingression of the sporozoites into the liver tissue. A conserved region C-terminal part of the protein binds to heparan sulfate proteoglycans associated with the surface membrane of hepatocytes [33, 97]. CSP is the most abundant molecule on the parasite surface: the number of molecules per parasite was estimated to around 10^6 [82].

The overall structure of CSP is conserved throughout different *Plasmodium* species: the protein is anchored to the cell surface by a GPI-anchor and composed of three major domains, an N-terminal, a C-terminal domain and a central repeat domain (Figure 2A). However, different *Plasmodium* species have different CSP protein sequences, with major differences in the sequence of the central repeating amino acids. In *Pf* the repeat region of CSP consists of around 40 repeat units of a NANP amino acid sequence followed in proximity to the N-terminus by several alternated repeat sequences including NVDP cadences. In field isolates of *Pf* strains from an endemic region, different

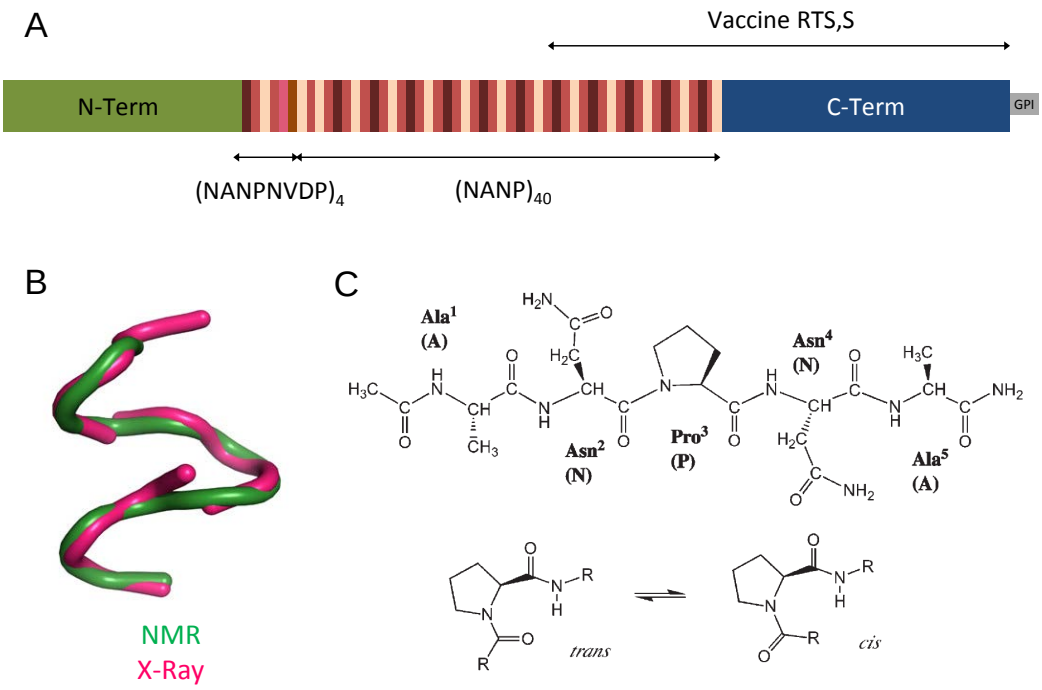


Figure 2. Structure of *Pf*CSP. A: *Pf* CSP is composed of an N-terminal, a C-terminal domain and a central repeat region composed mainly of NANP repeats but also few NVDP motifs in proximity to the N-terminus. It is anchored in the membrane with a GPI anchor. The stretch used in the vaccine RTS,S AS01 is indicated. B: NMR and X-ray studies have shown that single NANP repeats adopt a loop structure [14, 36, 79]. C: The NANP-repeat sequence can adopt type-I, type-II folding sequences due to rotamers of proline [36].

lengths of the repeat region and repeat numbers have been reported [15]. The repeat seems to evolve by concerted evolution, which means that repeat arrays can expand or contract by internal duplications or deletions [15]. Although a fully resolved structure of the CSP molecule does not exist, several studies describe structural characteristics of the smaller *Pf*CSP subregions like the C-terminus or single repeat-motifs [14, 36]. Summarizing these findings and using computational structure modelling, *Pf*CSP has been described to exhibit a flexible, rod-like structure with a high proportion of disordered protein stretches [79]. The *Pf*CSP component of the vaccine RTS,S is composed by the C-terminus and a stretch of NANP repeats.

The central repeat region of *Pf*CSP is very particular in its sequence and structure because it is composed by a long stretch of so-called tandem repeats, a characteristic which is seldom observed in proteins. By the use of NMR [14] or crystallography [36], it has been shown that the NANP-repeat adopts a loop structure [79] (Figure 2B). This secondary structure is facilitated by the prolines that can adopt both cis- and trans-conformations allowing the formation of type-I and type-II beta turns that are stabilized by hydrogen bonds within the repeat [36] (Figure 2C).

5.2 Humoral immunity

5.2.1 B cells generate specific antibodies in response to a foreign antigen

The adaptive immune system has evolved to efficiently and sustainably protect complex organisms from invading pathogens. The humoral immune system is the part of the adaptive immune system that generates antibodies as effector molecules. B cells are the cellular agents of humoral immunity responsible for antibody production. They develop in the bone marrow where they undergo multiple steps of differentiation and pass immune checkpoints that eliminate cells with potential auto-reactivity. After development every mature naive B cell carries a unique B cell receptor (BCR) on its surface, which is encoded by immunoglobulin loci that partially recombine during development.

Figure 3 depicts the evolution of B cells during an infection. Mature naive B cells that are able to recognize the foreign antigen with their BCR, are activated via a BCR-induced signalling cascade, proliferate, and migrate to secondary lymphoid organs such as lymphnodes or spleen to finally enter germinal center reactions (Figure 3A and B). In these germinal centers the foreign antigens are presented to B cells, which can internalize them if their BCR has a sufficient affinity to the antigen. The antigen is processed by the B cell and presented to follicular helper T cells that provide additional activation signals to those B cells that successfully presented the antigen (Figure 3D). Activated B cells proliferate and undergo clonal expansion. In addition, somatic hypermutation is used to introduce small alterations into the DNA sequence of the immunoglobulin gene and eventually change the affinity of the BCR to the antigen, a process called BCR diversification (Figure 3C). The enzyme primarily responsible for somatic hypermutation is called activation induced deaminase (AID), and raises mutation rates up to 10^{-3} mutations per base pair and division [12]. Antigen

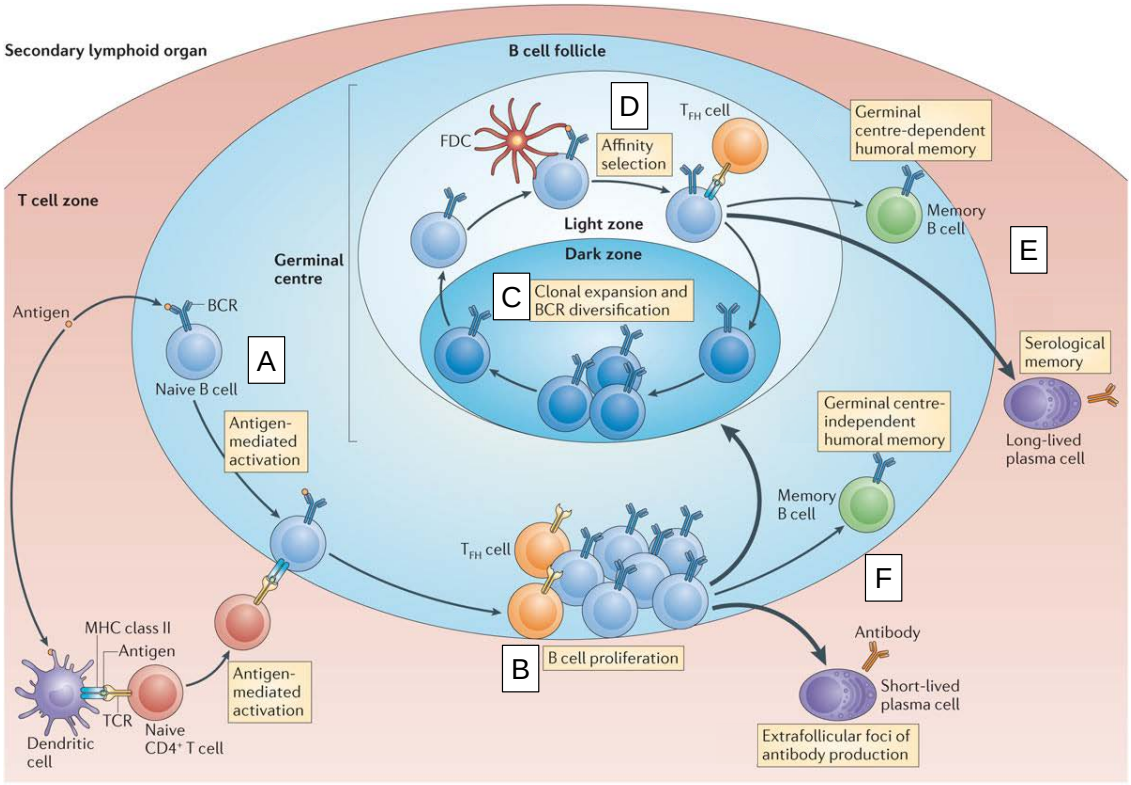


Figure 3. B cell activation and differentiation during the course of an immune reaction. Mature naive B cells are activated by binding of antigen to their B cell receptor and help of cognate CD4 T cells (A) and start to proliferate inside B cell zones of secondary lymphoid organs (B). Some of the activated B cells enter germinal centres where they undergo clonal expansion and BCR diversification by somatic hypermutation (C). In affinity selection single B cells compete for binding of antigen presented by follicular dendritic cells (FDC) and T cell help (D). When exiting germinal centers, affinity matured B cells can either differentiate into memory B cells or long-lived plasma cells (E). It has also been reported that some B cells differentiate into non-germinal center dependent memory B cells and plasma cells directly after first activation (F). Adapted from [55]

specific clones inside germinal centres thus compete for antigen binding and T cell help over several rounds of division and affinity maturation [109].

A B cell exiting the germinal center reaction can differentiate into either memory B cell or plasmablast (Figure 3E). Memory B cells are generally long-lived cells that keep their BCR surface expression, can be reactivated during a second infection and ensure a faster and more efficient secondary response upon reinfection with the same pathogen. Plasmablasts start to secrete antibodies instead of expressing the immunoglobulin genes as BCRs, a mechanism mediated by alternative splicing of the Ig transcript. They migrate to the bone marrow where they can differentiate into antibody-secreting plasma cells that provide antibodies to combat the current infection. Plasma cells can also maintain antibody titers to prevent secondary infections. The differentiation into either plasmablast or memory B cell is affinity and time dependent, with memory B cells produced in early germinal centers and higher affine plasmablasts in later stages [96, 119]. Plasmablasts are found in the peripheral blood approximately 7-10 days after an infection [17]. An atypical population of memory cells, characterized by their lack of the typical memory marker surface expression, has been described in chronic infections like HIV or malaria. Their exact origin and function remains unclear and studies draw diverging conclusions regarding their activation status, functionality and antibody secretion [63, 66, 80, 100].

In addition to the germinal center dependent generation of memory and plasmablasts, B cells have also been described to be able to differentiate into memory and plasmablasts directly after first antigen encounter without entering germinal centres [55] (Figure 3F).

5.2.2 Immunoglobulin structure diversity generation

Immunoglobulins, or antibodies, are Y-shaped molecules composed of two heavy (H) and two light (L) chains held together by disulphide bonds (Figure 4A). Each one of these chains is composed of a variable (V) and a constant domain (C). Different functions of the antibody molecule can be attributed to the different structural domains: the constant domain is responsible for antibody effector functions such as complement activation or opsonization, whereas the variable domain is responsible for antigen specificity. The two heavy constant domains are connected via disulphide bonds at the level of the antibody's hinge region. The hinge region is responsible for the flexibility of the two antibody arms. Different effector functions and constant domain structures or flexibilities are covered by the different isotypes. In humans these isotypes are IgD, IgM, IgG1, IgG2, IgG3, IgG4, IgE, IgA1 and IgA2. During the course of an immune response, the gene encoding the constant domain of a BCR can be replaced by a recombination process (class switch recombination) that also involves the above mentioned germinal center dependent enzyme AID.

The tips of the variable heavy and light chain domains compose the paratope of an antibody, i.e. the site where the antigen usually is bound. Subregions of the variable region which build the antigen-binding surface are called complementarity determining regions (CDR), whereas the protein

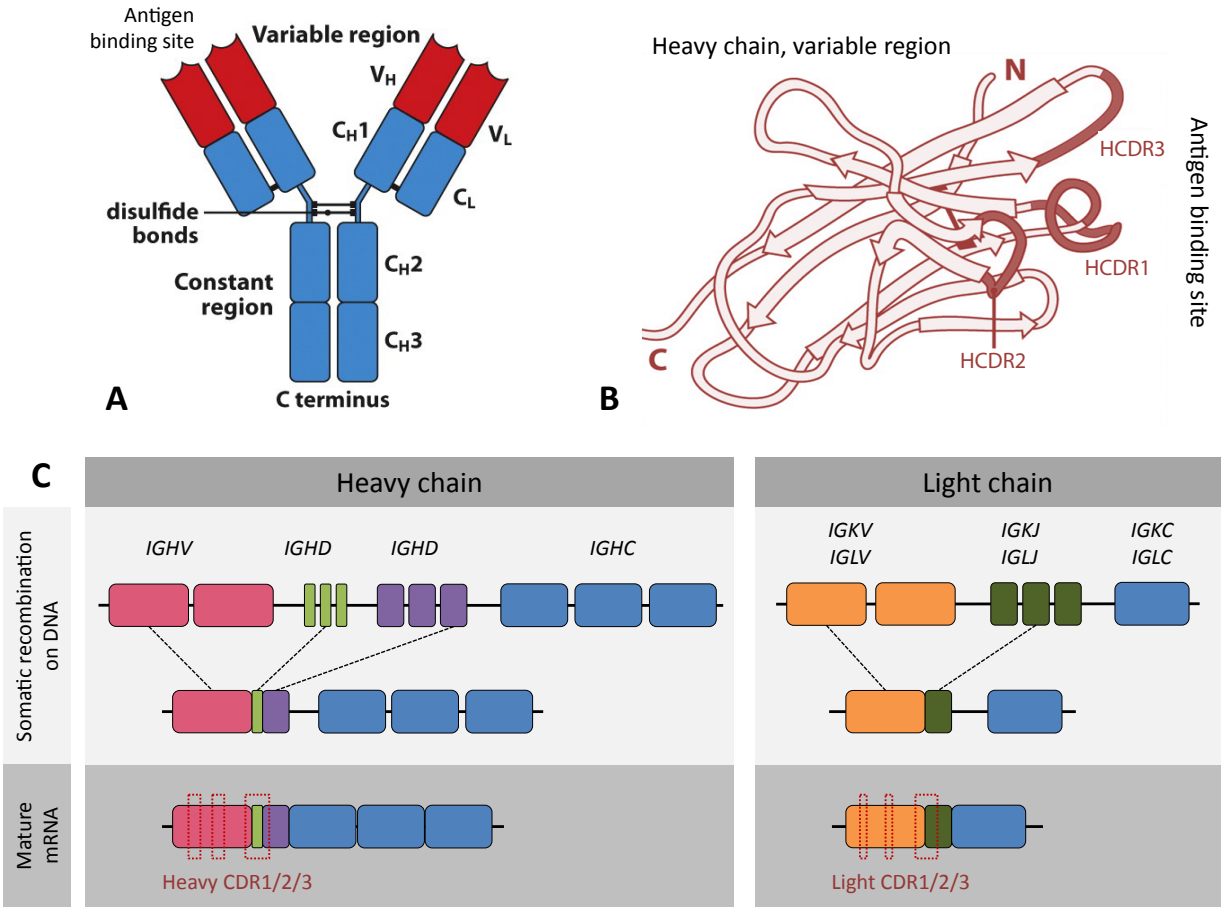


Figure 4. Structure and diversity of antibody molecules. A: Structure of an antibody molecule with heavy (H) and light (L) chains. Each chain is composed of a variable (red) and a constant (blue) region. The variable region forms the antigen binding site. The heavy chains are linked via disulphide in the hinge region. B: Protein structure of a variable heavy chain shown as cartoon. The complementarity determining regions (HCDR) define the antigen binding site. Framework regions (FWR) are shown in light pink. C: V(D)J recombination of heavy and light chain loci. V, D and J segments are recombined to encode the variable region. CDR1 and CDR2 are encoded by the IGHV gene, whereas CDR3 is located at the V(D)J recombination site and is thus the most variable. In the mature mRNA the exons encoding the constant region are assembled by splicing. Modified from [67].

Table 2. Number of Ig genes according to the IMGT database [56]

Locus	Chromosomal localization	V	D	J	C
IGH	14q32.33	123-129	27	9	11
IGK	2p11.2	76	0	5	1
IGL	22q11.2	73-74	0	7-11	7-11

stretches distal from the paratope are called framework regions (FWR) (Figure 4B). The location of these regions is usually defined based on analysis of surfaces of known antibody structures, as for example in the Kabat scheme [123]. During B cell development the genomic regions encoding the variable domain of the BCR are recombined (Figure 4C). For the heavy chain a VDJ recombination takes place, which means that one *IGHV*, one *IGHD* and one *IGHJ* segment are chosen from a pool of germline segments and joined by a process that involves deletion and addition of so called palindromic (P) and non-template (N) nucleotides. The light chain can be either encoded by the kappa or the lambda locus where a VJ recombination takes place. Table 2 summarizes the numbers of heavy, kappa and lambda germline genes present in humans.

The aleatory recombination of gene segments, the addition and deletion of nucleotides in the junction as well as the combination of heavy and light chain are responsible for the diversity of the human naive B cell repertoire. The actual diversity of the complete human naive heavy chain repertoire has been estimated to 10^{18} possible sequences, i.e. much larger than the approximately 10^{12} B cells present in an individual [28].

In addition to the intrinsic generation probabilities of single sequences, the B cell repertoire is also influenced by the genetic background of every individual. In studies describing the heterogeneity of the loci encoding human immunoglobulin variable chains, it has indeed been found that the variable germline genes are not only subject to single nucleotide polymorphisms or allelic variants, but that a great proportion of the diversity is also due to structural polymorphisms such as deletions or duplications [114].

5.2.3 Dynamics of antibody antigen recognition

During antigen binding of BCRs or antibodies, the immunoglobulin paratope formed by three heavy and three light CDR regions interacts with the epitope of an antigen. The molecular forces determining the affinity of the interaction are electrostatic forces, hydrogen bonds, hydrophobic interactions and van-der-Waals forces. Different types of shape complementarity have been described as for example long HCDR3 loops intruding binding pockets in large antigen molecules [91], or binding grooves for antigen peptides [85]. It has also been demonstrated that specific residues like tyrosine are more often involved in antigen binding than other amino acids [84, 86].

The affinity of an antibody (Ab) interaction with an antigen (Ag) interaction is defined as the equilibrium constant of the reaction

$$Ab + Ag \rightleftarrows AbAg.$$

The affinity constant of the reaction at steady state is then

$$K_D = \frac{K_{off}}{K_{on}} = \frac{[Ag][Ab]}{[AbAg]},$$

where K_{on} and K_{off} are the association and dissociation rates of the reaction which can also be determined individually by affinity measurement in a surface plasmon resonance setup for example.

Since antibody molecules have at least two binding sites (more for multimeric forms of IgM e.g.), avidity, which represents the overall binding rate of an antibody complex to an antigen, plays a major role. On the B cell surface, binding of BCRs to antigen and the thereby induced cross-linking induces B cell activation. It has been reported that if the nature of the antigen allows efficient cross-linking of B cell receptors, the B cell can get activated even without T cell help. These antigens are called T cell independent antigens and are particularly bacterial polysaccharides and other repetitive antigens.

5.2.4 Public antibody repertoires

Despite the enormous heterogeneity of recombination sequences, several studies have described the development of antibodies with convergent amino acid sequence features in multiple individuals after the exposure to common immune stimuli [35, 48, 76, 98, 105]. Even when sequencing complete antigen-experienced memory repertoires of individuals, public antibodies were identified, although with a extremely low frequency of 1 in 35,000 antibodies [24]. These common or public antibodies are of biological interest, because antibodies with similar sequence also are likely to have similar functions and binding behaviours. The successful identification and characterization of public antibodies could for example help to diagnose certain diseases or also to determine the effectiveness of a specific vaccination.

However, the difficulty in this approach lies in the successful clustering method that not only identifies antibodies with similar sequences but also only antibodies where similar sequences coincide with similar functionality. Antibodies are quite similar already due to the process that generates them at first. Various approaches have been used to identify public antibodies: complete HCDR3 identity, partial HCDR3 identity, similarity of only H chains or paired H and K chains. Datasets used in these analyses often stem from very different sources, including bulk repertoire sequencing that does neither allow an assessment of the complete antibody sequence nor the mapping of immunolgobulin sequences to single cells. The explanatory power of such investigations, however, is strongly

influenced by the quality and resolution of the utilized datasets. Very little work was done so far on testing the function of the identified public antibodies.

5.3 Open questions

5.3.1 Public antibody repertoires against the pre-erythrocytic parasite stages

The present study seeks to characterize the public antibody repertoire generated in humans after natural exposure and controlled human malaria infection. Previous research on antibodies sequenced and cloned from the controlled human malaria infection of *Pf* naive European donors which demonstrated pre-erythrocytic immunity [64] identified a considerable number of protective anti-*Pf*CSP antibodies [69]. The majority of these antibodies was of the IgM isotype and specific to the central NANP-repeat region of the *Pf*CSP protein. This study also identified common sequence features of the anti-*Pf*CSP antibodies, as for example usage of *IGHV3-33* and KCDR3 regions of 8 amino acids. However it remains unclear whether these sequence features also represent a common binding mechanism of antibodies to *Pf*CSP or whether they are enriched for reasons that are independent of antigen binding.

Up to date there are no high-throughput datasets on naturally acquired antibodies against pre-erythrocytic antigens, mainly because these antibodies are very rare and thus difficult to isolate. Previous studies on naturally acquired antibodies only analysed small numbers of these and did not identify sequence features that characterized the *Pf*CSP specific repertoire [73, 104]. It thus remains unclear to what extent naturally acquired antibodies are similar to the ones generated during controlled human malaria infection.

5.3.2 Anti-*Pf*CSP antibodies

Antibodies binding *Pf*CSP can be in general directed against the C-terminus, the N-terminus, or the central repeat region. However, this study focusses on antibodies with affinity to the NANP-repeat, because they are the ones most often observed. Screening of anti-*Pf*CSP antibodies against a library of synthetic repeat peptides has shown that these antibodies can be grouped according to the repeat cadences that they preferentially bind. The succession of N and P amino acids seems to be the most important cadence [122]. Indeed, recently resolved crystal structures have shown that antibodies can bind different small structural epitopes within the repetitive NANP-sequence [75, 104].

However, it is still not clearly described what sequence features mediate the binding of antibodies to the different repeat epitopes. Interestingly, for the antibodies from the controlled human malaria infection it was found that high affinities to the repeat can be reached even with very low numbers of somatic hypermutations [69]. This indicates that high quality germline encoded antibodies against *Pf* sporozoites exist. Similarly, two naturally acquired repeat binding antibodies with higher number

of somatic hypermutations have been shown to utilize mainly germline encoded residues for antigen binding [104]. Especially for the design of new vaccination strategies, it is important to understand how to efficiently target the germline precursors of high-quality B cell responses.

Another phenomenon that is still not well described is the impact of the *Pf*CSP repeat structure on the affinity maturation of the antibodies. In the antibodies isolated during the CHMI trial it was observed that affinity maturation plays a minor role for acquisition of high-affine antibodies and that clonal selection is more relevant. A mathematical model identified the complexity of the *Pf*CSP molecule as a reason for this phenomenon [69]. Moreover, biophysical investigations on murine and human antibodies directed against *Pf*CSP have shown that the repetitive structure of the protein favours multimeric binding of antibodies [30, 75].

6 Objectives

The overall goal of this thesis was to describe the molecular features of the public human antibody repertoire directed against the major sporozoite surface antigen *Pf* CSP. More specifically, the study had the following aims:

- Aim 1: Compare the *Pf* CSP-specific B cell repertoires in malaria naive and malaria experienced donors after controlled human malaria infection.
- Aim 2: Test to what extent the *Pf* CSP-specific public B cell repertoire can be defined using sequence data.
- Aim 3: Molecularly characterize several antibodies from this public repertoire and define the germline encoded sequence characteristics that determine their binding.

7 Material and Methods

7.1 Controlled human malaria infections and sample preparation

7.1.1 Tübingen controlled human malaria infection (TuCHMI)

Parts of the repertoire data and antibodies analysed in this thesis were generated earlier using samples from the TuCHMI controlled human malaria infection conducted by Benjamin Mordmüller and colleagues in Tübingen, Germany [64, 69]. In this study *Plasmodium falciparum* naive donors under chloroquine chemoprophylaxis treatment were infected three times with 51,200 live sporozoites with 28-day intervals. After another ten weeks the donors were challenged. The relevant time points for the present work are blood samples collected at day 7 after parasite injection. If not specified otherwise, the methods utilized for acquisition of the Tübingen data were the same as the ones described in the following sections.

The study was approved by the ethics committee of the medical faculty and the university clinics of the University of Tübingen and strictly adhered to Good Clinical Practice and the principles of the Declaration of Helsinki. The clinical trial from which the samples were obtained was registered under https://clinicaltrials.gov/ct2/show/NCT02115516 and number 2013-003900-38 in the EudraCT database and carried out under FDA IND 15862 and with approval of the Paul-Ehrlich-Institute [64].

7.1.2 Lambaréné controlled human malaria infection (LaCHMI)

The LaCHMI controlled human malaria infection was conducted in the Centre de Recherches Médicales de Lambaréné (CERMEL) at the Albert Schweitzer Hospital in Lambaréné, Gabon [57]. The study was approved by the Gabonese national ethics committee (Comité National d'Ethique de la Recherche), and conducted under a US FDA Investigational New Drug application (IND). The study followed the principles of the Declaration of Helsinki in its 6th revision as well as the "International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use – Good Clinical Practice (ICH-GCP)" guidelines. The study is registered with ClinicalTrials.gov, number NCT02237586 [57].

20 presumably *Pf* experienced Gabonese and 5 *Pf* naive donors aged between 18 and 29 years were pretreated with antibiotics 5 days prior to the study begin to eliminate all possibly remaining parasites from previous infections. At day 1, the donors were injected with 3,200 live *Pf* sporozoites intravenously. Starting on day 5, blood smear parasitemia screens were performed every day. 2 mL of peripheral blood was collected every second day by venous puncture. 20 mL of blood for B cell receptor sequencing were collected on day 7, 13, 19 or 21 and 28.

Donors were determined positive when at least 2 parasites were found on a slide by at least two independent slide readers (day of parasitemia, DoP). European donors were treated with anti-*Pf*

chemotherapy (Coartem) immediately after first positive thick blood smear, whereas Gabonese donors were treated when developing symptoms or reaching a threshold parasitemia of 1000 parasites per µL (day of Drug, DoD). Donors, who did not develop parasitemia, were treated on day 28, to eliminate all possibly remaining parasites.

7.1.3 Isolation of PBMCs

Peripheral blood was collected in Heparin tubes and centrifuged at 800g for 20 min. The serum was collected and the remaining blood components diluted 1:1 with RPMI medium. Peripheral blood mononuclear cells (PBMCs) were isolated using Percoll gradient density centrifugation at 800g for 20 min. The PBMC layer was isolated using a Pasteur pipette and washed twice in RPMI. $5x10^6$ PBMCs per mL and per vial were resuspended in 10% DMSO in FCS and frozen using a CoolCell cell freezing device and a -80 °C freezer. After 24 h the tubes were transferred to liquid nitrogen. Frozen PBMC samples where shipped to Berlin, Germany using a dry shipper container cooled with liquid nitrogen. Serum samples were shipped on conventional ice at 4 °C and refrozen at -80 °C when arriving in Berlin.

7.1.4 *IGHV* genotyping

Genomic DNA was extracted from whole blood (TuCHMI donors), or from sorted mature naive (MN) B cells (LaCHMI donors). *IGHV3* genes were amplified using a set of barcoded 5' primers and a mix of J-segment specific 3' primers (appendix, section 13.3.4). The PCR mix composition is shown in table 3, the PCR program was the same as for the 1st PCR amplification described in table 8. The amplicons were pooled and prepared for sequencing using a PCR-free library-prep kit (Illumina). Sequencing was performed on a Illumina MiSeq instrument using a 300-300bp paired end sequencing protocol. The DKFZ sequencing core facility offered the sequencing service and a basic sequencing quality control. Sequencing reads were then assembled using PandaSeq ([60], parameters: t = 0.8, minimum_length = 320, maximum_length = 550) and assigned to the donors by barcode identification. The resulting sequences were annotated using IgBLAST [126] and an in-house human genomic immunoglobulin database and pre-filtered on "IGHV3-3*"-like gene usage applying regular expression filtering. Every sequence was classified into the gene groups *IGHV3-33*, *IGHV3-30* or *IGHV3-3-30* depicted in Figure 17A using exact sequence matching of HCDR2 on nucleotide level.

7.1.5 Flow cytometric quantification of B cell populations

During the LaCHMI study, every second day 1-2 million isolated PBMCs were stained for flow cytometric analysis. PBMCs were centrifuged at 800g for 3 minutes and the pellets resuspended in

Table 3. PCR reaction mix for IGHV genotyping

PCR component	amount [µL]
PCR grade water	38.95
10x PCR Buffer	5
dNTPs [25 mM each]	0.4
Fw Primer [50 µM]	0.2
Rv Primer [50 µM]	0.2
HotStar *Taq* polymerase	0.25
Genomic DNA (2 ng/µL)	5

a primary staining mix composed of the following antibodies in 2% FCS in PBS (FACS buffer): CD19-PE-Cy7 (1:20), CD20-APC-H7 (1:10), CD21-PE (1:20), CD38-APC (1:20), CD27-FITC (1:5). After 30 minutes of incubation, cells were washed with FACS buffer and stained with the live/dead marker 7-AAD (1:400 in FACS buffer). After 10 minutes of incubation, the cells were again washed, resuspended in FACS Buffer and analysed on a Guava easy Cyte Cytometer. The data was analysed using FlowJo v10.0.

7.2 Antibody repertoire sequencing

7.2.1 Fluorescence activated cell analysis and single-cell sorting

PMBCs were thawed and washed in RPMI medium and pelleted at 500g for 5 min before the staining procedure. All sample preparation, staining steps and sorting was performed at 4 °C. NF54 *Pf*CSP was a kind gift of Dr. Kim Lee Sim (Sanaria Inc., Rockville, Maryland, USA) and was chemically coupled to Alexa647 according to labelling kit the manufacturer's protocol (Thermo Fisher Scientific Inc.). PBMC pelets were first resuspended and stained with the Alexa647-labelled *Pf*CSP in 2% FCS in PBS (FACS buffer) for 30 min.

Cells were washed with 1 mL FACS buffer, pelleted as described above and stained for 45 minutes using with the following antibodies in FACS buffer: CD27-PE (1:10), CD38-FITC (1:10), CD21-PE-Cy7 (1:20), IgG-APC-H7 (1:20), IgD-BV510 (1:20), CD138-BV421 (1:20), CD19-BV786 (1:10), CD20-BV711 (1:20). Cells were again washed and stained for 10 min with 7-aminoactinomycin D (7AAD, 1:400) as a live-dead marker. Finally cells were washed, pelleted and resuspended in FACS buffer and analysed or sorted on a BD FACS ARIA II cell sorter with a cooled plate stage.

Cell doublets were excluded based on FSC-H/W and SSC-H/W. Single cells were sorted into 384-well plates containing 2 µL of lysis/RHP buffer. The plates were immediately sealed, frozen on dry ice and stored at -80 °C. The lysis/RHP buffer consisted of the following components as shown in table 4: the reducing agent DTT for the disintegration of secondary RNA structures, NP-40 for cell lysis, RHP to initiate the transcription process and the RNAse inhibitor RNAsin.

Table 4. Sort reaction mixture

Sort mix component	stock concentration	amount [µL]
water	-	1.4813
PBS	10x	0.0500
DTT	100 mM	0.1000
NP-40	10%	0.1375
RHP	300 ng/µL	0.1375
RNAsin	40 U/µL	0.0938
total		2.0000

Table 5. Reverse transcription reaction mixture

RT mix component	stock concentration	amount [µL]
water	-	0.6375
RT-buffer	5x	0.8000
DTT	100 mM	0.3000
dNTPs	25 mM each	0.1375
RNAsin	40 U/µL	0.0563
SuperScript III/IV	200 U/µL	0.0688
total		2.0000

7.2.2 Ig gene amplification from single B cells

Sorted 384-well plates were shortly thawed on ice and then incubated at 68 °C for 60 sec to break secondary RNA structures. 2 µL of the RT mix for cDNA synthesis were added. The RT mix was composed of the following reagents as shown in table 5: RT enzyme buffer, DTT, dNTPs, RNAsin and the reverse transcriptase SuperScript III. The reaction was carried out on a Eppendorf Matercycler with the following steps: denaturation of RNA (5 min, 42 °C), annealing of hexamers (10 min, 25 °C), reverse transcription (60 min, 50 °C), reaction stop (5 min, 94 °C).

The cDNA synthesis product was used as template for a first Ig gene specific PCR. The product of this reaction was then used for a second gene specific PCR with barcoded primer sets. The PCR composition and cycler protocol is described in Tables 6 - 8. The gene specific parts of both primer sets are shown in the appendix, section 13.3.1. All pipetting steps for the RT-PCR were carried out on a Tecan Evo200 automation platform as described in [16].

7.2.3 Sequencing of Ig amplicons

In order to enable high throughput sequencing on an Illumina platform, sequencing adaptors have to be ligated to the amplicons. Illumina TruSeq PCR-Free LT Kit was used for adaptor ligation according to the manufacturer's protocol. The amount of amplicons with adaptors ligated to both

Table 6. First PCR reaction mixture with Ig specific primers

1st PCR reagent	stock concentration	amount [µL]
water	-	7.7900
buffer	10x	1.0000
5' primer mix	50 µM	0.0325
3' primer mix	50 µM	0.0325
dNTPs	25 mM each	0.1000
HotStar *Taq* polymerase	5 U/µL	0.0450
template		1.0000
total		10.0000

Table 7. Second PCR reaction mixture with Ig specific primers

2nd PCR reagent	stock concentration	amount [µL]
water	-	7.7900
buffer	10x	1.0000
5' primer mix	50 µM	0.0325
3' primer mix	50 µM	0.0325
dNTPs	25 mM each	0.1000
HotStar *Taq* polymerase	5 U/µL	0.0450
template		1.0000
total		10.0000

Table 8. PCR programs for Ig gene amplification

Step	Duration	Heavy	Kappa	Lambda	Cycles
Initial denaturation	10 min	95 °C	95 °C	95 °C	1
Denaturation	30 sec	95 °C	95 °C	95 °C	
Annealing	30 sec	58 °C	58 °C	60 °C	50
Elongation	55 sec (1st PCR), 45 sec (2nd PCR)	72 °C	72 °C	72 °C	
Final elongation	10 min	72 °C	72 °C	72 °C	1

ends was measured using the KAPA Quant qPCR kit on a Roche 480 light cycler instrument. Amplicons were then pooled in appropriate ratios and sequenced on a Illumina MiSeq platform using a 300-300 bp paired end sequencing protocol in the DKFZ Sequencing Core Facility. 30-40 % of PhiX DNA was added to the sample before sequencing to reduce biases due to non-heterogeneous base distribution.

7.3 Computational sequence analysis

7.3.1 Sequence annotation using sciReptor

Paired end reads from the Illumina run were assembled using PandaSeq [60]. The parameters set during the pandaseq run were: t = 0.8, minimum_length = 320, maximum_length = 550. Assembled reads were analysed using the sciReptor pipeline for analysis of single-cell sequencing data ([47], github.com/b-cell-immunology/sciReptor). Briefly, single reads were attributed to their respective well barcodes identified by RazerS analysis [118]. V and J segments as well as somatic hypermutations, constant segment usage and CDR/FWR regions were annotated. The flow cytometric index data and meta-information was associated to every single cell using the R flowCore package [29]. The data was stored in a MySQL database framework using a MariaDB server. Subsequent sequence analysis and plotting was done using python 2.7, especially the matplotlib [45], numpy [106].

7.3.2 Sequence similarity and repertoire clustering

In order to characterize repertoire similarity, pairwise sequence distance between all observed sequences was calculated independently for heavy, kappa and lambda repertoire. For this purpose sequences were aligned using the Biopython alignment algorithm pairwise2.align.globalxx [20]. As distance measure the hamming distance was calculated with gaps and mismatches counting 1. The distance was normalized to the length of the alignment. For clustering of the resulting pairwise distance matrices python networkx package was used [38]. Plotting of circular cluster diagrams was done using circos [54].

7.3.3 Background model for somatic hypermutation

The background model for somatic hypermutation was taken from the shazam package with standard settings [37, 125]. Sets of computationally mutated IGHV3-33, IGHV3-23 and IGKV1-5 sequences were generated in fashion to have similar mutation levels as the observed datasets. Computed and observed sequences were translated using Biopython [20] and the amino acid frequencies compared at defined positions.

7.4 Recombinant antibody expression

7.4.1 Cloning of immunoglobulin genes

After sequence analysis, specific antibodies were selected for recombinant expression, in order to test their antigen binding.

Preparation of heat competent *E. coli*

E. coli DH10B were grown in 500 mL of lysogeny broth (LB) medium to an OD600 of 0.5. After an incubation of 30 min on ice cells were centrifuged for 10 min at 3200g and 4 °C and the pellet was resuspended in 70 mL of ice-cold 0.1 M $CaCl_2$. The bacterial cells were centrifuged, resuspended in 50 mL 0.1 M $CaCl_2$ solution supplemented with 15% glycerol and snap-frozen in liquid nitrogen. 50 µL aliquots of the competent bacteria were stored at -80 °C. The transformation efficiency (competency) was determined by test transformation vector in amounts ranging between 1 and 10^{-5} ng of DNA.

Expression vectors

Three vectors were used for expression of antibody heavy and light chains: AbVec2.0-IGHG1, AbVec1.1-IGKC and AbVec1.1-IGLC2-XhoI [103]. DNA preparation of these vectors was performed as described in "Transformation into heat competent bacteria" and "Vector preparation". Maps of the three expression vectors are depicted in the Appendix, Section 13.4.

Specific amplification of immunoglobulin genes

In order to clone the heavy and light chain amplicons into the respective expression vector, another PCR was carried out using primers which introduce restriction sites at the 5' and 3' end of the heavy and light chain amplicons. Primer sequences can be found in the appendix, section 13.3.2. SalI, BsiWI and XhoI were used as 3' restriction sites for heavy, kappa and lambda chains respectively. The 5' restriction site was an AgeI site for all three chains. In the specific PCR, 3.5 µL of the 1st PCR mixed with 7.5 µL of water were used as template. The reaction set-up is shown in table 9. The PCR was run with the thermocycler program for the secondary PCR of the heavy chain (Table 8). Amplification was verified by gel electrophoresis. PCR products were purified using the NucleoSpin 96 PCR clean-up kit (Macherey-Nagel GmbH).

Table 9. PCR reaction for amplification of Ig genes for cloning

Specific PCR reagent	stock concentration	amount [µL]
water	-	21.42
buffer	10x	4.00
5' primer mix	3.3 µM	2.00
3' primer mix	3.3 µM	2.00
dNTPs	25 mM each	0.40
HotStar *Taq* polymerase	5 U/µL	0.18
total		26.0000

Restriction endonuclease digestion of vector and inserts

The amplicons of the heavy and lambda chain of the Ig genes obtained in the specific PCR and the purified Igγ1 and λ vectors were double-digested with the respective enzymes (heavy: AgeI-HF/SalI-HF, lambda: AgeI-HF/XhoI, New England Biolabs) for 2 h at 37 °C. As different incubation temperatures for the enzymes AgeI and BsiWI were required, a 2-step digest was carried out for the amplicons of the kappa chain and the Igκ vector, first digesting with AgeI-HF for 2 h at 37 °C and then for 2 h at 55 °C after adding BsiWI. All digests were carried out according to the enzyme manufacturer's protocol. The digested PCR products were purified using the NucleoSpin 96 PCR Clean-Up kit from Macherey-Nagel GmbH. The digested vectors were separated by gel electrophoresis using a 1% agarose gel. The linearized vectors were excised and the DNA purified according to the manufacturer's protocol using the NucleoSpin Gel and PCR Clean-up kit (Macherery-Nagel GmbH).

Ligation

The digested PCR products of the heavy, kappa and lambda chains were ligated into the respective expression vectors using T4 DNA Ligase. Ligation was carried out at 16 °C overnight.

Transformation into heat competent bacteria

For transformation 10 µL of heat competent bacteria were transferred to a microcentrifuge tube on ice. 3 µL of ligation product (or 0.4 µL of vector DNA in the case of retransformation) were added and the mixture incubated on ice for 30 minutes. A heat shock at 42 °C was performed for 45 seconds. The bacteria were cooled down again on ice for 2 minutes before adding 100 µL of LB-medium. The cultures were incubated for 40 minutes at 37 °C and 650 rpm and then plated on LB-agar plates containing 100 µg/mL ampicillin.

Table 10. PCR reaction mixture for insert check

Insert check PCR reagent	stock concentration	amount [µL]
water	-	20.975
buffer	10x	2.5
5' primer mix	50 µM	0.2
3' primer mix	50 µM	0.2
dNTPs	25 mM each	0.125
HotStar *Taq* polymerase	-	1.00
total		25.0000

Screening of bacterial colonies by PCR

Transformed bacterial colonies were screened for the correct insertion of the heavy, kappa or lambda variable insert. For this purpose, a colony PCR was performed where bacterial colonies were directly added to the PCR master mix described in Table 10.

The 5' primer Absense, which binds to a vector sequence upstream of the inserted PCR product, was used for all three vectors. Primers specific for the individual constant regions of human Igγ1, Igκ and Igλ were used as 3' primers: hIGHG-084-Rv for Igγ1, hIGKC-172-Rv for Igκ and hIGLC-057-Rv for Igλ. PCR products were analyzed on a 2% agarose gel for their expected sizes of 650 bp for Igγ1, 700 bp for Igκ1 and 590 bp for Igλ. To exclude PCR-induced point mutations or frame-shift occurring through the insertion of the PCR products into the vectors, the PCR products were Sanger sequenced with Absense primer. The sequences were compared to the sequences initially obtained during repertoire sequencing. A list of the primers can be found in the appendix, section 13.3.3.

Vector DNA preparation

For miniprep, 6 mL of terrific broth (TB) medium with 75 µg/mL ampicillin were inoculated with a bacterial colony. For maxiprep, 500 mL of LB Medium with 75 µg/mL ampicillin were inoculated. In both cases DNA preparation was performed according to the manufacturer protocol using plasmid purification kits (Macherery-Nagel GmbH). DNA concentration and purity were assessed using a NanoQuant plate in a Tecan M1000 Pro plate reader.

7.4.2 Site directed mutagenesis

Site directed mutagenesis on the antibody encoding plasmids was performed using the Q5 site directed mutagenesis kit (Quiagen). The primers are summarized in the Appendix, Section 13.3.5. The successful introduction of mutations was verified by sequencing the resulting vector using the Absense primer.

Table 11. Volumes, cell numbers and cell culture conditions for transfection

Container	Volume cells	Cell density (M per mL)	Shaking (rpm)	DNA per chain (µg)
48 well plate	1 mL	0.8	180	0.5
50 mL falcon	10 mL	1.5	180	15
600 mL flask	100 mL	1.5	130	150

7.4.3 Expression of antibodies in a mammalian cell line

HEK-293F cell culture

FreeStyle293-F cells (Invitrogen) were cultured in a volume of 20 mL FreeStyle 293 Expression Medium at 37 °C and 5% CO_2 on a shaker plate at 180 rpm. Every 2-3 days cells were split to a density of 5x10^5 cells per mL.

Transfection of HEK-293F cells using polyethylenimin (PEI)

One day prior to transfection, cells were seeded in the appropriate culture container described in Table 11. On the day of transfection, for 10 mL and 100 mL cultures, the DNA of heavy and light chain plasmids was added to the culture and the mixture incubated for 5 minutes. 150 µL or 1.5 mL of PEI solution (0.6 g/L) were then added respectively and the cells returned to culture conditions. For the 48 well plate, DNA for heavy and light chain was mixed on a sterile 96 well plate with conical bottom. The PEI solution was diluted 1:8 in FreeStyle 293 medium and 40 µL/well of this dilution added to the DNA plate. The DNA/PEI mixture was incubated for 5 minutes and transferred to the cells using a multichannel pipette. The cells were cultured for 24 h and on the next day, one culture volume of Excell medium was added.

Purification of expressed antibodies

Recombinant antibodies were purified from the cell culture supernatant using Protein G beads for FastFlow (GE Healthcare). Antibodies were bound to the appropriate amount of beads, calculated based on the binding capacity, by incubating over night at 4 °C. Beads were harvested by centrifugation and then loaded on empty columns and washed with two column volumes of PBS. The antibodies were eluted using 0.1 M glycine at pH 3 and re-buffered to pH 7 using 1 M Tris at pH 9. Purified antibodies were dialysed against PBS using Slide-A-Lyzer Dialysis devices (Thermo Scientific) according to the manufacturer's protocol.

7.5 Antibody testing

7.5.1 Antibody concentration measurement

The antibody concentration of the cell culture supernatant or of the purified antibodies was measured by enzyme-linked immunosorbent assay (ELISA). 384-well high-binding transparent plates were coated overnight at 4 °C with 15 µL/well PBS containing 1.66 µL/mL of the goat anti-human IgG Fc-fragment capturing antibody.

Plates were washed three times with deionized water using a Tecan Plate Washer. 20 µL of blocking buffer (PBS with 0.05% Tween and 1mM EDTA) were added to the plate and incubated for 1 h at room temperature to block unspecific binding sites. Again the plate was washed, and 12.5 µL/well of diluted supernatant or purified antibodies were added. For every measured antibody a 1:2.5 dilution series of 8 steps was set up starting with an initial dilution between 1:10 and 1:2000 according to the estimated stock concentration. As a standard, human IgG1 antibody standard of known concentration was used in a total of 16 dilution steps starting at a concentration of 3 µg/mL.

After 1.5 h of incubation the plate was washed. 15 µL of a horseradish peroxidase (HRP) conjugated goat anti-human IgG diluted to 0.83 µL/mL in blocking buffer were added to each well and incubated for 1 h. The ELISA was then washed three times and developed by adding 25 µL/well of ABTS 1-step solution (Roche) previously mixed with 1 µL/mL H_2O_2. Absorbance was measured at 405 nm after 5 and 10 minutes on an M1000Pro plate reader (Tecan) using the Magellan software and the concentration of each sample was calculated from the standard curve.

7.5.2 Antigens

Four different antigens were used for antibody binding assays in this work: ΔN-CSP is a truncated version of *Pf*CSP with only amino acids 123-411, i.e. with an N-terminal deletion [101]; $NANP_{10}$ and $NANP_5$ with 10 or 5 NANP repeat units (Alpha Diagnostic International); $NANP_3$ with 3 NANP repeat units (PSL GmbH, Heidelberg).

7.5.3 Antigen ELISA

High-binding 384 well plates were coated with different antigens overnight at 4 °C. ΔN-CSP was used at 10 ng/well, $NANP_{10}$ at 25 ng/well, $NANP_5$ at 50 ng/well, $NANP_3$ at 82.5 ng/well with 25 µL/well. The plates were washed and blocked with 1% BSA in PBS for 1 h at room temperature. The sample antibody was diluted in 8 steps of 1:4 starting at 4 µg/mL, transferred to the coated plate and incubated with the antigen for 1.5 h at room temperature. Bound IgG was detected using an anti-human IgG-HRP and the corresponding detection kit in the same way as for the concentration ELISA described in section 7.5.1. The OD value at 405 nm was measured after 5 and

10 minutes on a Tecan M1000 instrument using the iControl software. The humanized version of a mouse anti-CSP antibody 2A10 [104] and an isotype control mGO53 [112] were used as positive and negative control respectively.

7.5.4 Surface plasmon resonance

Surface plasmon resonance measurements were performed on a BIACORE T200 (GE healthcare) instrument docked with a series S sensor chip CM5 (GE healthcare). 10 mM HEPES buffer with 150 mM NaCl, 0.02 % Tween 20 and 0.05 % BSA at pH 7.4 was used as a running buffer. Anti-human IgG antibodies were immobilized on the chip using an amine-coupling based human antibody capture kit (GE Healthcare). In a first step, the concentration of all sample antibodies was determined using a human IgG1 kappa standard from myeloma plasma. In the second step, equal concentrations of sample antibody and isotype control were captured in the sample and the reference flow cell, respectively. Running buffer was injected for 20 min at a rate of 10 µL/min in order to stabilize the flow cells. $NANP_5$ and $NANP_3$ at 0, 0.015, 0.09, 0.55, 3.3, 20 µM in running buffer was injected at a rate of 30 µL/min. After each sample antibody the flow cells were regenerated with 3M MgCl2. The data were fit using a 1:1 binding model in the BIACORE T200 softwareV2.0.

7.5.5 *Plasmodium falciparum* traversal assay

Cells from a human hepatocyte cell line (HC-04, MRA-975, deposited by Jetsumon Sattabongkot) were seeded at $6x10^4$ cells/well in a 96-well plate (source) and incubated for 24 h at 37 °C and 5% CO_2. 7,500 *Pf* sporozoites obtained from mosquito salivary glands were incubated with different concentrations of monoclonal antibodies for 30 min. The mixture was added to the cells in presence of 0.5 mg/mL dextran-rhodamine. Untreated sporozoites were used as positive control and cells incubated with only dextran-rhodamine were used to determine the experimental signal background. After sporozoite addition the plate was centrifuged at 3000 rpm for 10 min without brakes and incubated for 2 h at 37 °C and 5% CO2. The cells were washed with PBS three times, trypsinized and resuspended in 10% FCS in PBS. They were then centrifuged as 3600 rpm for 5 min and the pelet resuspended in 1% PFA in PBS. The percentage of dextran positive, i.e. traversed cells was measured using a BD LSR II flow cytometer. The background signal was subtracted from all measurements and the percentage of traversal inhibition calculated at the ratio of the maximal traversal rate obtained from untreated sporozoites.

7.6 Antibody structure modelling

Computational models of the antibody variable region structure as well as docking models of antibody to antigen structures were obtained using the Rosetta Antibody software together with a recently

published protocol [120]. For each antibody 3050 models were generated and the overall score used to select the best results. Ensemble docking using the Rosetta snugdock algorithm was performed on these 10 models and the crystal structure of an $NANP_5$ repeat bound to an IGHV3-33/IGKV1-5 antibody (kindly provided by J.-P. Julien). For structural clustering, the 50 best models of each antibody were selected and pooled for the complete set of antibodies that was supposed to be analysed. The structures were aligned using the PyRosetta [18] alignment function and the CDR regions extracted. GOSSIP, a length independent structure alignment algorithm, was used to cluster the CDR regions [52]. Structural visualization was done using PyMol (V1.8.2.2, DeLano Scientific LCC, Schrödinger). The antibody modelling approach was developed during a research internship in the laboratory of Prof. William Schief in the course of an EMBO short term fellowship.

8 Results

8.1 Selection of *Pf*-experienced donors from CHMI cohort and sorting of B cell subsets

As part of the investigation of public antibody repertoires in pre-erythrocytic immune responses, one major aim of the present work was to investigate whether malaria experienced donors from an endemic area raised or carried similar antibodies as the previously studied European donors after CHMI. Since previous work showed that naturally acquired antibody responses to pre-erythrocytic antigens are rare, the present work focused on samples collected during a controlled human malaria infection in malaria experienced donors. It was speculated that due to an active immune response during the study, humoral immune responses against the sporozoites and antigen specific B cells may be easier to detect and isolate.

PBMC samples were collected from a CHMI of donors from a malaria endemic region in Lambaréné, Gabon (LaCHMI, [57]). In a first step the immunity against the parasite infection as well as the humoral immune response parameters of the LaCHMI participants were assessed in order to choose donors with signs of protective anti-*Pf* CSP antibody response for further analysis.

8.1.1 Donors show differences in development of parasitemia and malaria symptoms

LaCHMI participants were infected intravenously with 3,200 sporozoites and monitored throughout the course of 28 days. Figure 5A shows the overall outcome of the study in terms of blood stage parasitemia and malaria symptom development. The day of parasitemia is the day on which the first blood smear was found to be positive, i.e. at least two blood stage parasites were identified on one slide. The day of drug is the day on which participants were administered the anti-*Pf* chemotherapeutic lumefantrine-arethmeter because they had developed malaria symptoms. All five European control subjects (*Pf* naive, NA) developed blood stage parasitemia between day 12 and day 14, as expected for a *Pf* infection and confirming the validity of the infection protocol. Out of the 20 Gabonese donors, which were considered to be pre-exposed (PE) because of living in an endemic area, 12 donors developed blood stage parasitemia between day 13 and day 25. This prolongation of the prepatency period indicates a higher level of immunity against *Pf* sporozoite infection or development of blood stage parasitemia. Some of these donors (e.g. L1-003) even showed a delayed or absent development of symptoms, a phenomenon known as asymptomatic parasitemia. 8 Gabonese donors did not develop blood stage parasitemia at all during the course of 28 days after infection and were thus considered to be protected against the infection due to naturally acquired pre-existing immunity.

A

Donor ID	Day of parasitemia	Day of drug	Group
L1-014	**12**	**12**	**NA**
L1-015	**12**	**12**	**NA**
L1-025	**12**	**12**	**NA**
L1-024	**13**	**13**	**NA**
L1-001	**14**	**14**	**NA**
L1-006	13	16	PE
L1-008	14	16	PE
L1-017	15	19	PE
L1-020	16	16	PE
L1-019	16	19	PE
L1-026	17	28	PE
L1-005	18	18	PE
L1-010	19	23	PE
L1-003	19	28	PE
L1-028	23	28	PE
L1-023	24	24	PE
L1-018	25	28	PE
L1-002	n.d.	28	PE
L1-007	n.d.	28	PE
L1-009	n.d.	28	PE
L1-011	**n.d.**	**28**	**PE**
L1-013	n.d.	28	PE
L1-016	n.d.	28	PE
L1-021	**n.d.**	**28**	**PE**
L1-022	**n.d.**	**28**	**PE**

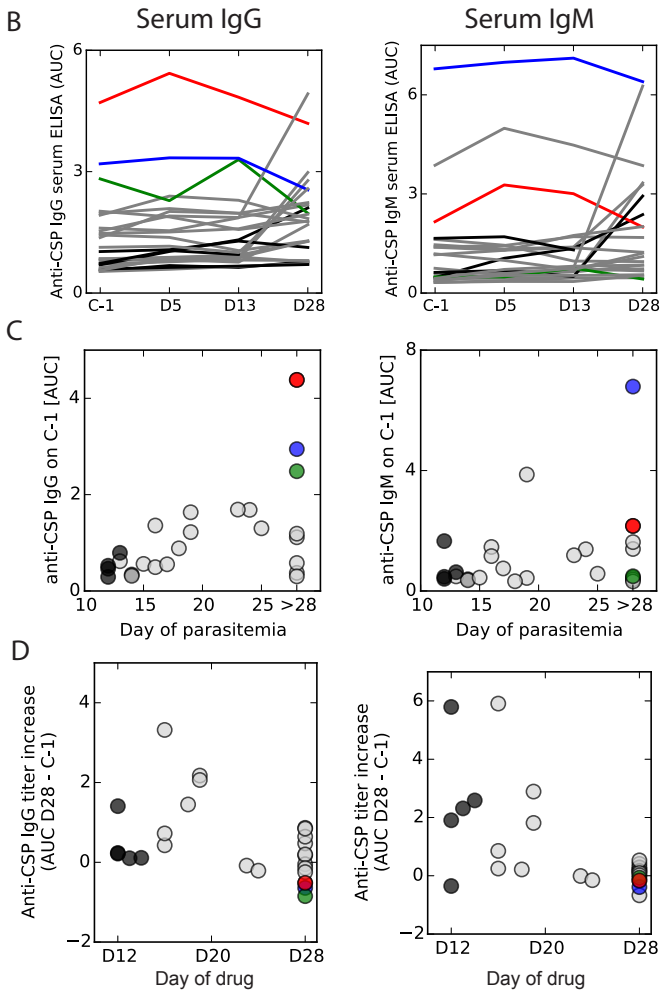


Figure 5. Selection of LaCHMI donors based on immunological criteria. A: Day of blood stage parasitema and day of drug treatment for LaCHMI donors. NA: *Pf* naive, PE: *Pf* pre-exposed. B: Anti-*Pf*CSP IgG and IgM serum titers at days C-1 (one day before infection), D5, D13 and D28. C: Anti-*Pf*CSP IgG and IgM serum antibody levels (AUC) on day C-1 as a function of the day of blood stage parasitemia. D: Difference in anti-*Pf*CSP IgG and IgM serum antibody levels (AUC) between Day C-1 and D28 as a function of the day of drug treatment. Donors selected for further analysis are highlighted in red, blue and green. European donors are shown in black, Gabonese donors in grey.

8.1.2 Some donors show anti-*Pf*CSP serum titers

The focus of the present work is on antibodies directed against the repeat region of the sporozoite surface protein *Pf*CSP. In order to choose suitable donors based on the humoral immune response to *Pf*CSP, anti-*Pf*CSP IgG and IgM serum titers were measured at different time points before and after the sporozoite injection. Figure 5B shows the area under curve (AUC) measured in serum ELISA against *Pf*CSP one day before infection (C-1) and 5, 13 and 28 days after infection (D5, D13 and D28). Most of the donors had no measurable or only a weak serum antibody response against *Pf*CSP during the course of the study. A clear correlation between anti-*Pf*CSP serum titers before injection and the protection level could not be established (Figure 5C). Notably, many donors were protected against the infection even without measurable *Pf*CSP serum antibody at C-1. Some of the donors raise their anti-*Pf*CSP serum antibody levels during the course of the study (Figure 5B and D), indicating an active immune response against the sporozoites. Interestingly, the donors with higher serum reactivity at D28 than on C-1 are also the ones that were treated early on in the infection (Figure 5D). In donors that do not develop blood stage parasitemia or symptoms the anti-*Pf*CSP reactivity even sometimes drops during the course of the study (Figure 5D).

The highlighted donors L1-011, L1-021 and L1-022 were chosen for repertoire sequencing because they were protected against the infection and showed the highest anti-*Pf*CSP IgG titers on C-1. This choice is based on the assumption that common protective anti-*Pf*CSP antibodies were most likely found in donors with pre-existing humoral immune memory against *Pf*CSP.

8.1.3 Heterogeneity in the B cell population dynamics

In order to investigate the dynamics of the B cell response in different donors, the proportions of different B cell populations as plasmablasts (PB), classical memory B cells (CM) and atypical memory B cells (AtM) were monitored during the course of the LaCHMI study. Figure 24 (appendix) depicts the gating strategy chosen to identify cells belonging to the different B cell sub-populations. The population frequencies are represented in section 13.1.1 (appendix). No correlation between protection, parasitemia and B cell dynamics could be established. The dynamics of the PB population were very heterogeneous, with some donors showing none, one or two peaks of high PB frequencies. The population frequency data was not used to select donors for repertoire sequencing.

8.1.4 Sorting and characterization of plasmablasts, atypical memory B cells and *Pf*CSP binding B cells

For the three selected donors L1-011, L1-021 and L1-022 single-cell sorting of different B cell populations was performed in order to sequence paired heavy and light immunoglobulin chains and characterize the *Pf*CSP specific antibody repertoire. Fluorescently labelled full-length *Pf*CSP was

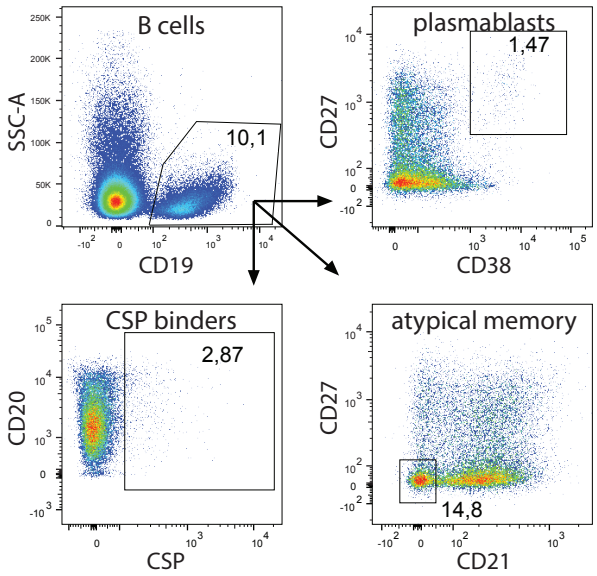


Figure 6. Gating strategy for isolation of CSP binders, plasmablasts and atypical memory B cells from CD19+ B cells. The B cell gate is pre-gated on single live lymphocytes using side and forward scatter as well as the live/dead stain 7-AAD. Data shown for donor L1-021 on day D7.

used as a bait to isolate *Pf*CSP binding B cells. In addition *Pf*CSP specific B cells are expected to be activated and differentiate into antibody secreting plasmablasts after the sporozoite injection. They can be found in the blood 7-10 days after first antigen encounter but since they loose BCR surface expression, they are inaccessible to identification with a fluorescently labelled antigen bait. In contrast the role of AtM B cells remains controverse, but it has nevertheless been shown that they can be specific to malaria antigens [34, 66]. In order to isolate potential *Pf*CSP specific B cells in the PB (CD19+, CD27+, CD38+) and AtM B cell population (CD19+, CD21-, CD27-), these two populations were specifically sorted at D7 and D13 after infection. The detailed gating strategy for donor L1-021 on D7 is depicted in Figure 6.

Figure 7A confirms that the AtM B cells share characteristics of previously observed atypical subsets, as notably elevated surface expression of CD19 and CD20 in comparison to CM B cells and partial IgG surface expression. Interestingly, the frequency of AtM B cells was higher in the *Pf*CSP binding B cell population than in the overall B cell population in most of the samples (Figure 7C). The AtM B cells enriched in the *Pf*CSP binding compartment share the same phenotype as the overall AtM population (Figure 7B).

The previous observations and results thus allowed to select three donors out of the Gabonese CHMI cohort based on signs of pre-existing humoral immunity to *Pf*CSP. Moreover analysis of the B cell populations used for single-cell sorting revealed that AtM B cells are enriched in the *Pf*CSP antigen specific B cell population.

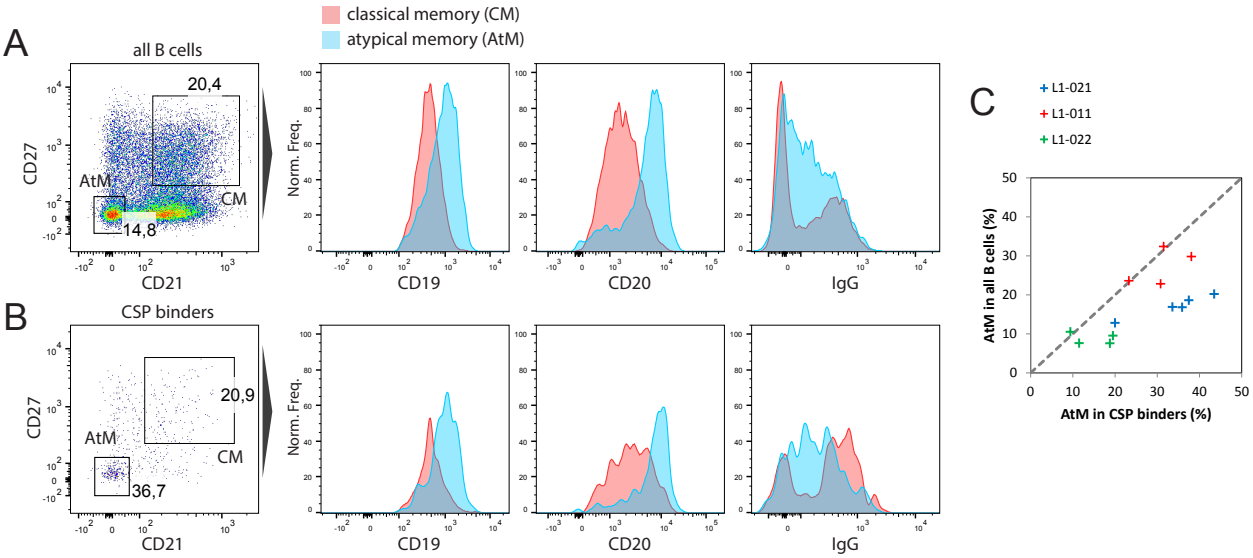


Figure 7. Enrichment of atypical memory B cells in the *Pf*CSP binding population. A: CD19, CD20 and IgG surface expression of atypical memory (AtM) and classical memory (CM) B cells pre-gated on all CD19+ B cells. B: CD19, CD20 and IgG surface expression of AtM and CM B cells pre-gated on *Pf*CSP binding B cells. C: Frequency of AtM in CD19+ B cells against AtM in *Pf*CSP binding B cells.

Table 12. Constant primers in Roche454 [68] setup and Illumina setup

Locus	PCR reaction	Roche454 constant primers	Illumina constant primers
IGKC	1st PCR	hIGKC-220-Rv	hIGKC-172-Rv
	2nd PCR	hIGKC-172-Rv	hIGKC-032-Rv
IGHA	1st PCR	hIGHA-140-Rv	hIGHA-111-Rv
	2nd PCR	hIGHA-111-Rv	hIGHA-076-Rv
IGHG	1st PCR	hIGHG-137-Rv	hIGHG-137-Rv
	2nd PCR	hIGHG-084-Rv	hIGHG-074-Rv

8.2 Sequencing and computational annotation of B cell repertoires

In order to be able to compare the *Pf*CSP specific antibody repertoires, the immunoglobulin transcripts of the sorted B cells needed to be amplified, sequenced and annotated. In general the previously published method [16] relies on a matrix of 384-well plates into which single cells are sorted well-wise. A high-throughput automation platform is used to perform massive-parallel reverse-transcription PCR reactions on single-cell level. The resulting amplicons are pooled for next generation sequencing. Due to the discontinuation of the Roche 454 sequencing platform, the established human transcript amplification method [68] had to be adapted to an Illumina sequencing platform. In addition a bioinformatics pipeline had to be established to analyse the sequencing results.

8.2.1 Generation of immunoglobulin amplicons for Illumina sequencing

The longest read lengths that can be sequenced and reliably assembled using the Illumina technology are around 550 bp and can be achieved using the Illumina MiSeq 300bp-300bp paired end sequencing protocol. To adapt the previous Ig transcript amplification method to Illumina sequencing, one major step was to develop new primer sets to reduce the length of the amplicons, which was previously up to 600 bp. Notably the 3' primers specific for constant regions *IGHG*, *IGHA* and *IGKC* had to be replaced, whereas the 5' primers specific for the variable region were not modified (Table 12). Figure 8 shows the comparison of amplification efficiencies between the old primer sets as published in [68] and the primers developed for the new sequencing approach. The amplification efficiency, i.e. the percentage of successfully amplified transcripts assessed by gel electrophoresis, could be maintained or sometimes even increased with the new primer sets.

Also, since Illumina sequencing requires unconventional Y-shaped sequencing adaptors, the new primer sets did not include the sequencing adaptors any more as opposed the previous approach [16]. The adaptors were instead ligated to the amplicons during library preparation using a Illumina TruSeq PCR-free library prep kit.

8.2.2 Development of a computational tool for single-cell level B cell repertoire analysis

To analyse the raw sequencing data provided by the Illumina MiSeq instrument a new computational analysis tool was developed under the name of "Single cell immune Repertoire tool-kit" (sciReptor, [47]). The role of the program is to map single reads to the wells they originated from by identifying paired tag sequences. It also annotates standard immunoglobulin characteristics as gene segment usage, mutational status, isotype and others. The data and annotations are stored in a relational database.

The first step in the computational analysis is to determine which immunoglobulin chain (heavy (H), kappa (K) or lambda (L)) is encoded by each sequence and to identify the forward and reverse tags

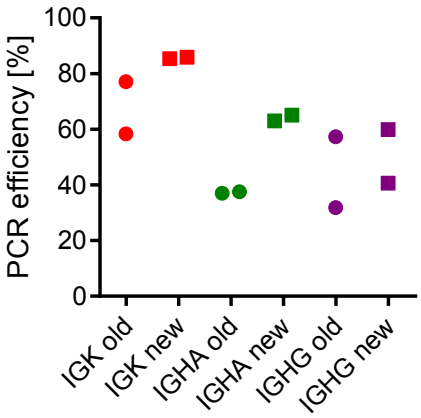


Figure 8. PCR efficiencies of the primer sets published in [68] compared to the newly developed primer sets for generation of shorter fragments, as shown in Table 12. Efficiency was tested on sorted Igκ+ memory B cells (IGK), sorted IgA+ memory B cells (IGHA) or sorted IgG+ memory B cells (IGHG). The amplification efficiency represents the percentage of successfully amplified transcripts (assessed by gel electrophoresis) within the total set of amplification reactions.

that are necessary to locate the well of origin of each sequence. IgBLAST [126] is used for locus annotation, whereas RazerS [118] is used for tag finding. The statistics of tag identification for one exemplary run are shown in Figure 9A. 807,256 raw sequencing reads went into the analysis of which 485 failed to be attributed to a specific locus. Between 208,000 and 315,000 reads were associated to the three loci and between 40 and 80% of these had a successful identification of both tags. The relatively low level of tag identification especially for the L locus is due to sequencing errors in the tag region or due to sequencing of amplicons that were amplified with primers originating from the first PCR reaction and thus not containing sequence tags.

After successful identification of the sequence tags, sciReptor assembles all reads belonging to a certain well and locus and builds consensus sequences using MUSCLE sequence alignment [27]. Since B cells sometimes have two immunoglobulin transcripts from independent recombination events, the algorithm first assembles the sequences mapped to the most abundant V-J combination to build a first consensus. In the following step, the second most common V-J combination is treated similarly. Figure 9B depicts the statistics of first and second consensus and their respective reads per well numbers (red and blue curve). The histogram shows the overall distribution of reads per well. A consensus is only built if more than 5 sequences are assigned to the well. The fact that the first consensus clearly has the majority of sequences per well confirms the validity of sequencing and consensus building approach.

Figure 9C finally represents the spatial distribution of sequencing success on the 12 384-well plates belonging to one matrix processed by the automation platform. Every black square on the matrix represents a successful identification of the corresponding sequence. In the case of the depicted experiment the success rate of sequencing the heavy and corresponding light chain was 54%. Notably,

the efficiencies obtained after computational sequence analysis are comparable to those measured by counting the PCR positive wells on an agarose gel, confirming that the computational approach is suited for the analysis of the single-cell repertoire sequencing data.

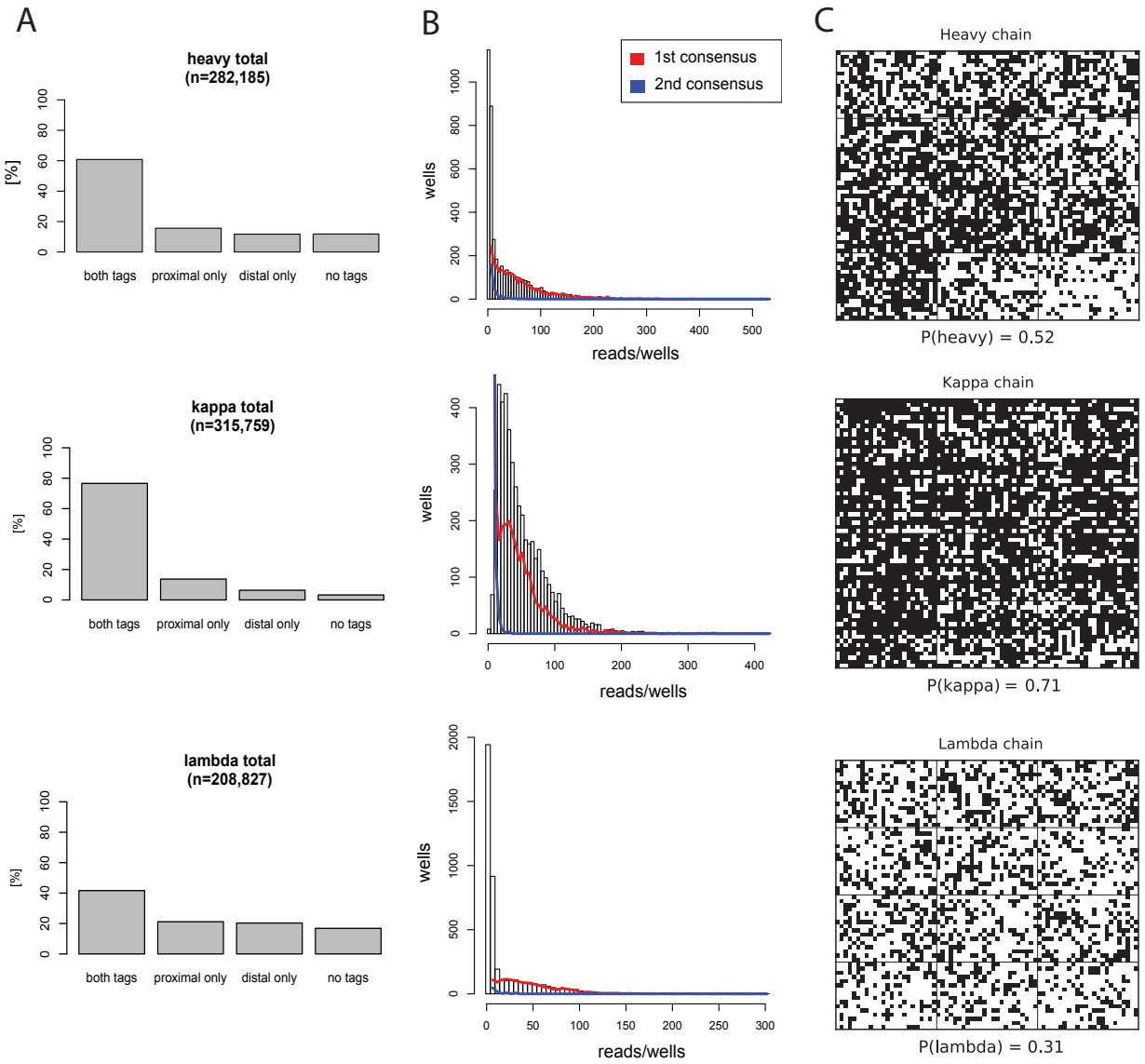


Figure 9. Quality control for the successive steps of the computational analysis of repertoire sequencing data. A: Statistics of identification of proximal and distal tags for H, K and L chains. B: Reads per well statistics for all sequences (histogram), sequences assigned to 1st consensi (red) and sequences assigned to 2nd consensi (blue). C: Spatial representation of the 12 384-well plates belonging to one processing matrix. Black squares indicate successful sequencing of the H, K or L immunoglobulin transcript.

8.3 Public repertoire of *Pf*CSP-reactive antibodies is restricted to European donors

Once the methodology for repertoire sequencing was established, the immunoglobulin transcripts of the different B cell populations from the Gabonese donors were sequenced and annotated. In order to identify clusters of public repertoire in the CSP-specific repertoire of Gabonese and European donors, the antibody sequences were compared regarding their amino acid sequence similarity. Pairwise full-length alignments were performed with sequences between the start of CDR1 and the end of the J segment. The sequence homology was calculated as the ratio of Hamming distance divided by the length of the alignment. A graph was built from the similarity matrix obtained from the pairwise alignments. In this graph, nodes correspond to single antibodies, i.e. paired heavy and light chain sequences. Edges have two weight attributes: sequence similarity on heavy chain level and sequence similarity on light chain level. Table 13 summarizes the two different data sets that went into the analysis.

8.3.1 Community clustering approach identifies small and diverse clusters of similarity

In order to identify public antibodies with similar sequence characteristics, the analysis was further focused on single groups of similar antibodies. To this purpose the graph generated during computation of similarity distributions was reduced to a set of subgraphs with an edge weight of at least 90% in heavy and light chain.

Figure 10 shows a selection of the biggest repertoire clusters identified using the clustering approach and grouped according to the utilized heavy chain variable gene. Every unit on the circles represents one antibody sequence belonging to the different Gabonese LaCHMI and European TuCHMI donors. Lines indicate a sequence similarity of more than 90% on amino acid level in heavy and light variable region. It can be observed that the clusters are small and diverse, utilizing a great variety of different V segments. Nevertheless an utilization of *IGHV* genes from the *IGHV3* and *IGHV4* family seems to be preferred, most importantly *IGHV3-23*, *IGHV3-33* and *IGHV4-59*. The public repertoire clusters are more often shared between European than between Gabonese donors.

Table 13. Dataset used for the analysis of repertoire similarities

Dataset	Description
TuCHMI	2600 paired H+L chain sequences Data from 5 European donors *Pf*CSP binders (non mature naive) and plasmablasts sorted from PBMCs [69]
LaCHMI	1012 paired H+L chain sequences Data pooled from 3 Gabonese donors *Pf*CSP binders, plasmablasts and atypical memory B cells sorted from PBMCs

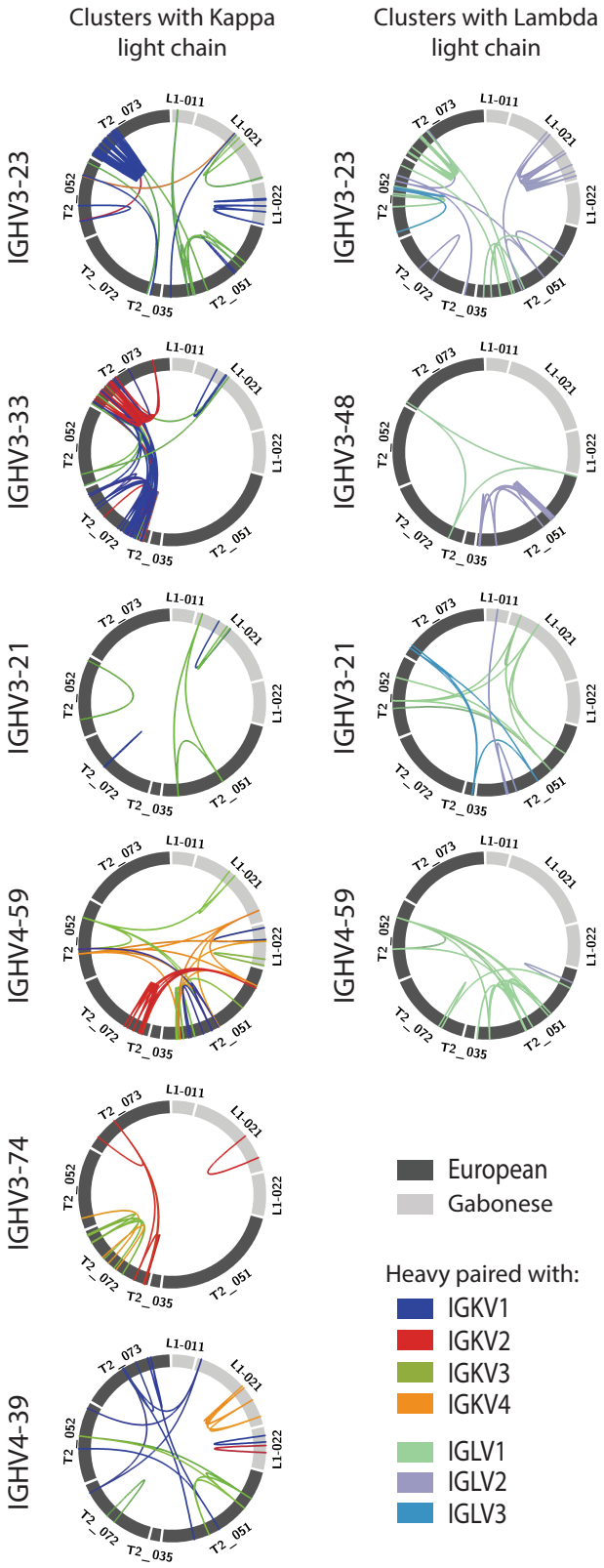


Figure 10. Selection of clusters of similarity with >90% amino acid homology in heavy and light chain. Every unit of the circle represents one antibody sequence, grouped by European (dark grey) and African (light grey) donors. Connecting lines represent an amino acid homology of >90% in heavy and light chain. Left column: Clusters utilizing Igκ light chains. Right column: Clusters utilizing Igλ light chains. Colours encode the family of the utilized light chain variable gene.

8.3.2 Antibody cloning based on sequence similarity identifies clusters of similar binding only in European repertoire

In order to investigate, whether the clusters of sequence similarity identified in the previous step also represent groups of functional similarity, the binding behaviour of selected members was assessed. *Pf*CSP reactivity data was either taken from a previous study [69] or assessed by cloning, expressing and *Pf*CSP ELISA reactivity assessment of additional antibodies. To focus on antibodies against *Pf*CSP, clusters where at least one member bound *Pf*CSP in FACS sorting were chosen. Another criteria for the cluster selection was that they were shared by two donors at least.

Figure 11 shows a representative selection of sequence similarity based reactivity prediction results. Within most of the clusters the tested antibodies show similar reactivities, i.e. either binding to ΔN-CSP in ELISA above a defined threshold (AUC >1) or not. However, antibodies from Gabonese donors clustering together with reactive antibodies from European donors, as in cluster 83, fail to show measurable *Pf*CSP reactivity. In general, the reactivity prediction approach works well for European donors but not for Gabonese donors. No antibodies from Gabonese donors could be identified that bound *Pf*CSP.

8.3.3 Antibody cloning based on sequence annotation fails in Gabonese repertoire

In order to exclude that the *Pf*CSP antibodies in the public Gabones repertoire are too highly mutated to have enough sequence similarity to the European germline like antibodies, predictive cloning was also performed based on more general sequence features. It has e.g. previously been shown that IGHV3-33 is enriched in the high-affinity *Pf*CSP binding antibodies as well as antibodies with a short KCDR3 of 8 amino acids. A set of 10 antibodies with either of these sequence features were thus cloned from the Gabonese donors, expressed and tested in antigen ELISA. None of them bound *Pf*CSP in ELISA.

For one specific antibody utilizing *IGHV3-33* and *IGKV1-5* with a KCDR3 region of 8 amino acids, it was investigated whether somatic hypermutations could have abolished *Pf*CSP binding. A molecular model of the antibody was generated and the positions of somatic hypermutations leading to amino acid exchanged investigated [78]. However even when reverting the relevant mutation at position H.58, no *Pf*CSP binding could be observed, indicating that the antibody even in it's germline configuration was not *Pf*CSP reactive.

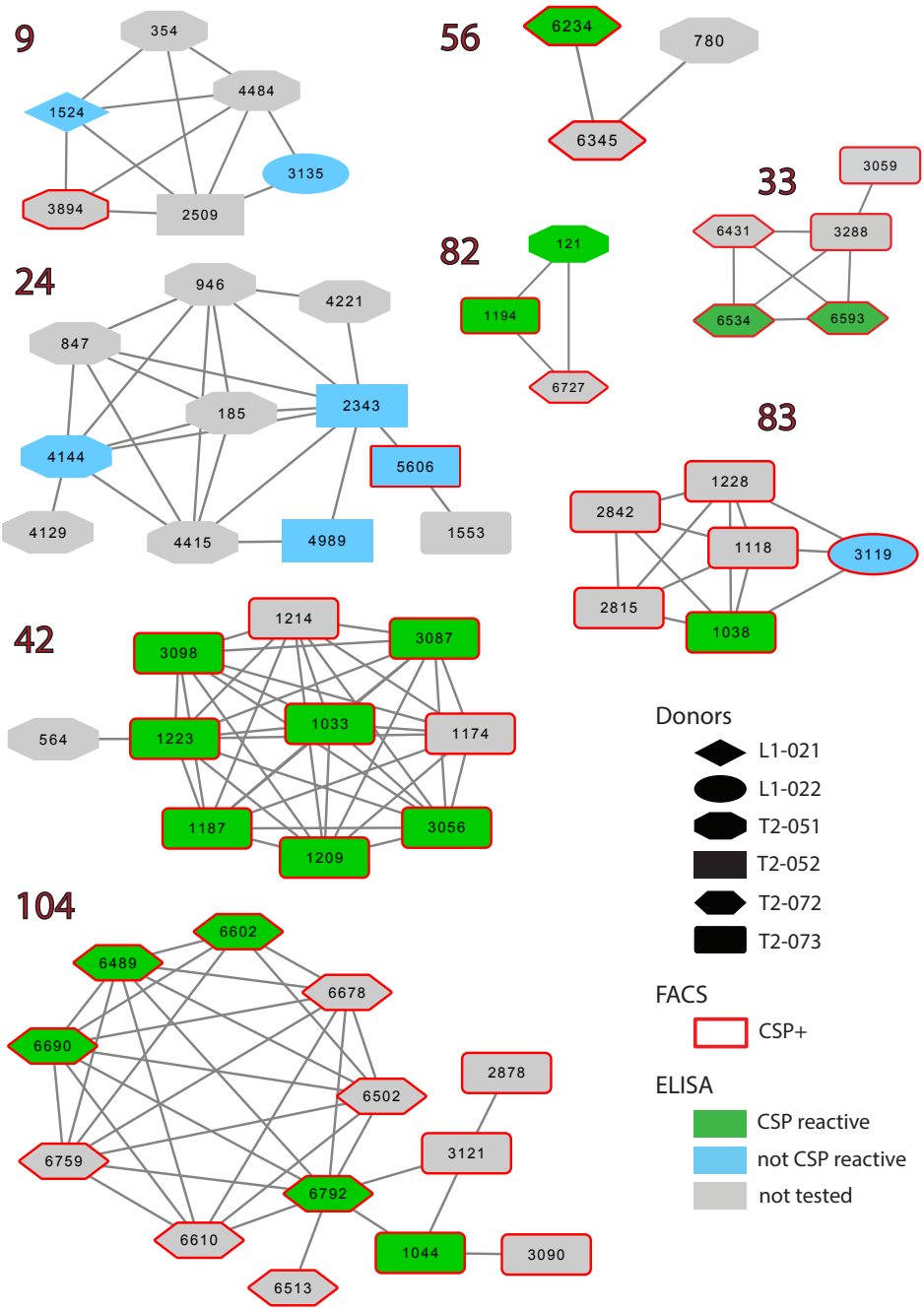


Figure 11. Predictive cloning based on identified clusters of similarity. Cluster number indicated in bold. Every node represents one antibody, i.e. paired heavy and light chain immunoglobulin sequences. The number inside the node represents the event identifier attributed by the analysis software. The shape of the node represents the donor from which the sequence was isolated. Grey lines connect antibodies with >90% amino acid similarity in heavy and light chain. Nodes corresponding to cells that bound CSP in FACS have a red border. The color fill indicates reactivity in ELISA (reactive: AUC >1).

8.3.4 Cloning of antibodies from Gabonese repertoire based on *Pf*CSP binding in FACS does not identify binders

Surprisingly none of the two previous predictive cloning approaches lead to identification of binding antibodies in the Gabonese donors. Since the donors nevertheless were selected based on their anti-*Pf*CSP titers and also showed *Pf*CSP specific memory B cells in FACS sorting, it was however assumed that CSP specific memory B cells were present in the isolated populations. A last selection of cells was thus made based on *Pf*CSP binding during FACS sorting, and the corresponding set of 41 antibodies cloned and tested in antigen ELISA. The CSP binding of the corresponding B cells in FACS as well as their isotype determined in sequencing are depicted in Figure 12A and B. The majority of the antibodies is of the IgM isotype, an observation that has previously also been made for the *Pf*CSP specific antibodies of European donors [69]. None of these antibodies showed *Pf*CSP-binding in ELISA12C, which could mean that their affinity to *Pf*CSP was not high enough or that the antibodies were not at all specific for *Pf*CSP.

In the Gabonese donors, none of the here described cloning attempts lead to successful identification of antibodies with measurable *Pf*CSP reactivity in ELISA. However, the similarity clustering as well as the prediction of *Pf*CSP binding based on sequence similarity gave satisfying results for the antibody repertoires of the European CHMI donors. Several clusters of highly *Pf*CSP reactive antibodies were identified and shared between different donors.

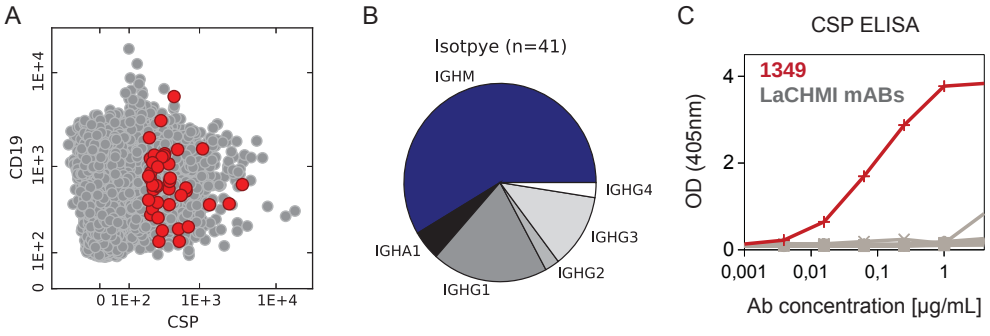


Figure 12. Antibodies cloned from B cells sorted using a fluorescently labelled *Pf*CSP bait. A: *Pf*CSP binding in FACS. Cells with transcripts selected for cloning are shown in red. All sorted cells are shown in grey. B: Isotype distribution of the transcripts selected for cloning. C: *Pf*CSP ELISA reactivity of cloned antibodies shown in grey. 1349 (red) is a positive control antibody presented in section 8.4.1.

8.4 Molecular characterisation of binding modes in public antibodies

The last part of the present work focuses on describing the molecular *Pf*CSP-repeat binding characteristics of common CSP-reactive antibodies isolated from the European CHMI donors. Understanding the molecular determinants of these antibodies, that were raised during repeated controlled sporozoite injection, could eventually lead to a better understanding of why such antibodies are not found after natural parasite exposure.

Two groups of antibodies with very low levels of somatic hypermutations ($<$5 amino acid exchanges from germline) and high *Pf*CSP reactivity were selected for molecular characterization and comparison:

- Antibodies with an Igλ chain utilizing *IGLV1-47* and a LCDR3 region of 11 amino acids. The heavy chains are encoded by different heavy chains, raising the question whether the binding mode is similar and whether the light chain plays a similar role in all antibodies.
- Antibodies with an Igκ chain utilizing *IGKV1-5* and a particularly short KCDR3 of 8 amino acids. Again different *IGHV* genes are used.

8.4.1 Comparison of CSP binding of antibodies utilizing *IGLV1-47* but different *IGHV3* variable genes

Assessment of antibody structure similarity by homology modelling

The first set of selected antibodies shares highly similar or even identical Igλ chains but utilize the heavy chain variables genes *IGHV3-48* or *IGHV3-23* (Figure 13A). In order to assess the structural similarity of the antigen binding site in these antibodies, computational homology modelling was performed for each antibody variable region. The 50 best structure models for every antibody obtained by Rosetta homology modelling were selected and pooled. Structural clustering of the CDR regions was performed based on a length-independent global structure alignment algorithm [52]. By applying this procedure to the pool of structure models, antibody models with similar CDR conformations were clustered, in order to detect a similar conformation spectrum of the antibodies that were modelled. Figure 13B depicts the size distribution of the resulting structural clusters. All antibodies encoded by *IGHV3-23* have models that cluster together in the biggest cluster (Nr. 0), indicating a convergence in structural modelling (Figure 13C). However the two antibodies utilizing *IGHV3-48* have fewer (1349) or no models (121) in cluster Nr. 0, but still cluster together in Cluster Nr. 1. The structural similarity within the *IGHV3-23* encoded antibodies thus seems to be greater than between the antibodies encoded by different *IGHV* genes.

A

Donor	IGHV	IGHJ	HCDR3	H SHM	IGLV	IGLJ	LCDR3	L SHM	CSP binding In FACS	Antibody name (if cloned)	ELISA Reacitvity
T2_072	*IGHV3-48*01*	*IGHJ4*02*	ARDSLGSSSWPNYFDY	1	*IGLV1-47*01*	*IGLJ2*01*	AAWDDSLSGVV	0	CSP+		
T2_073	*IGHV3-48*01*	*IGHJ4*02*	ARDAVRSSWPSYFDY	2	*IGLV1-47*01*	*IGLJ2*01*	AAWDDSLSGPV	1	CSP+	D07g1l1349	CSP binding
T2_051	*IGHV3-48*02*	*IGHJ4*02*	ARDGVGSSWPNYFDY	1	*IGLV1-47*01*	*IGLJ2*01*	AAWDDSLSGVV	0	CSP+	D0751x0121 *	CSP binding
T2_073	*IGHV3-23*01*	*IGHJ4*02*	AKDYSSSWPSYFDY	0	*IGLV1-47*01*	*IGLJ2*01*	AAWDDSLSGVV	0	CSP+	D0773x1015 *	CSP binding
T2_073	*IGHV3-23*01*	*IGHJ4*02*	AKDTDSNYPSYFDY	1	*IGLV1-47*01*	*IGLJ2*01*	AAWDDSLSGVV	0	CSP+	D0873x0860 *	CSP binding
T2_073	*IGHV3-23*01*	*IGHJ4*02*	AKDSSSGWPSYFDY	0	*IGLV1-47*01*	*IGLJ2*01*	AAWDDSLSGVV	0	CSP+	D0873x0871 *	CSP binding
T2_073	*IGHV3-23*01*	*IGHJ4*02*	AKDQSSGWPSYFDY	0	*IGLV1-47*01*	*IGLJ2*01*	AAWDDSLSGVV	0	CSP+	D0873x0829 *	CSP binding
T2_073	*IGHV3-23*01*	*IGHJ4*02*	AKDQSSGWPSYFDY	0	*IGLV1-47*01*	*IGLJ2*01*	AAWDDSLSGVV	0	CSP+	D0773x1199 *	CSP binding
T2_073	*IGHV3-23*01*	*IGHJ4*02*	AKDQSSGWPSYFDY	0	*IGLV1-47*01*	*IGLJ2*01*	AAWDDSLSGVV	0	CSP+	D0773x1340 *	CSP binding
T2_073	*IGHV3-23*01*	*IGHJ4*02*	AKDQSSGWPSYFDY	0	*IGLV1-47*01*	*IGLJ2*01*	AAWDDSLSGVV	0	CSP+	D0773x1006 *	CSP binding
T2_073	*IGHV3-23*01*	*IGHJ4*02*	AKDSSSGWPSYFDY	0	*IGLV1-47*01*	*IGLJ2*01*	AAWDDSLSGVV	0	CSP+	D0773x1001 *	CSP binding
T2_073	*IGHV3-23*01*	*IGHJ4*02*	AKDTDSNYPSYFDY	0	*IGLV1-47*01*	*IGLJ2*01*	AAWDDSLSGVV	0	CSP+	D0773x0982 *	CSP binding

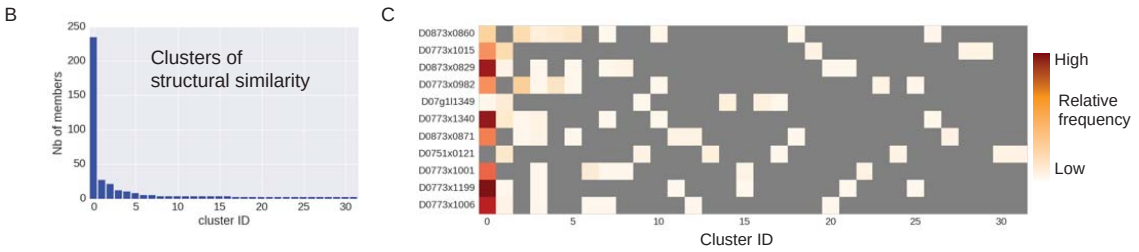


Figure 13. *IGLV1-47* utilizing antibodies and their structural similarity assessed by molecular modelling. A: Sequence properties of selected public antibodies utilizing *IGLV1-47*. Gene usage, somatic hypermutation counts as well as *Pf*CSP binding of corresponding B cells in FACS is indicated. All cloned antibodies were reactive to the repeat region of *Pf*CSP in ELISA. B: Size distribution of the clusters of structural similarity obtained by applying a global structure alignment algorithm to a pool of the 50 best structure models for every antibody. Parameters for GOSSIP clustering: alignment length threshold 0.95, maximal distance between two matching Cα 1.5 Å. C: Contribution of models from the different antibodies to the clusters of structural similarity. Cluster IDs are assigned according to the size of the cluster.

Both clusters show similar sensitivity for major changes in the lambda chain

Experiments introducing major changes in the lambda chain showed that the LCDR3 length as well as the sequence of the *IGLV1-47* segment play a role in antigen binding of the here probed antibodies (Figure 14). When replacing the light chain of antibody 1349 with a highly mutated *IGLV1-47* encoded light chain isolated from a Gabonese donor, repeat binding was lost, although the LCDR3 was still 11 amino acids long (_alt_l11, Figure 14A). Similarly, when pairing the heavy chain of 1349 with and unmutated *IGLV1-47* utilizing light chain with 12 amino acids, binding was lost (Figure 14A). Similar observations were made for antibodies 121, 982 and 1015 which are other representative antibodies utilizing *IGHV3-48* or *IGHV3-23* (Figure 14A).

Effect of affinity maturation and recombination sequence of the light chain is dependent on the heavy chain sequence

The loss of binding when paired with a highly mutated *IGLV1-47* encoded light chain already highlighted the importance of the exact amino acid sequence of the Igλ chain. Interestingly, one of the antibodies utilizing *IGHV3-48*, antibody 1349, shows an amino acid replacement in the Igλ chain at residues L.33 and an alternative LCDR3 region that differs in one amino acid due to recombination differences (Figure 15A). These residues are not part of the antibody surface as indicated by molecular modelling (data not shown) and thus most likely have a stabilizing effect on surrounding residues that directly interact with the antigen. In order to investigate what effect these replacements had on the other antibodies, a set of heavy chains utilizing either *IGHV3-48* or *IGHV3-23* were paired with the affinity matured light chain l1349 (Figure 15B). $NANP_{10}$ binding

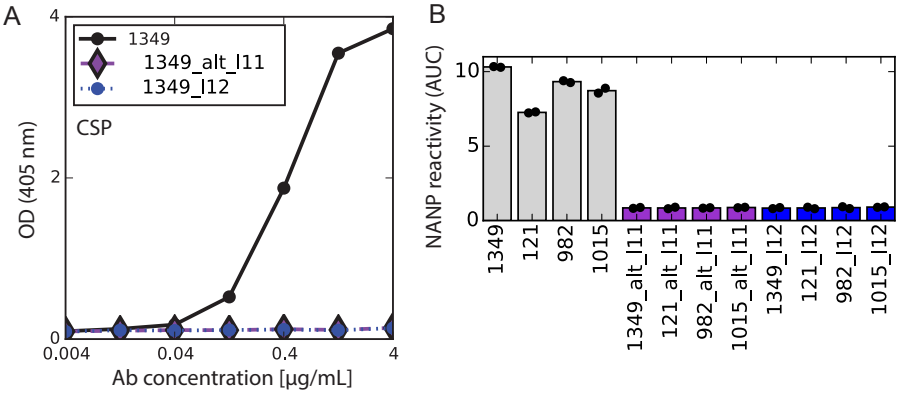


Figure 14. Importance of the LCDR3 length and sequence for *IGLV1-47* utilizing antibodies. A: $NANP_{10}$ antigen ELISA with 1349 wildtype and chimeric antibodies with 1349 heavy chain and mutated *IGLV1-47* utilizing light chain with 11 amino acid long LCDR3 (alt_l11) and *IGLV1-47* utilizing light chain with 12 amino acid long LCDR3 (l12). B: Loss of $NANP_{10}$ binding in antigen ELISA due to light chain replacement in 1349, 121, 982 and 1015 antibodies.

of 121, another antibody utilizing *IGHV3-48* was not affected or even slightly improved, whereas both antibodies utilizing *IGHV3-23* (982 and 1015) showed reduced binding. As expected, when replacing the mutated lambda chain of 1349 with a germline encoded lambda chain, the AUC was reduced. This shows that the effect of affinity maturation and recombination sequence in the 1349 light chain is dependent on the heavy chain and thus beneficial only for antibodies utilizing the same *IGHV* gene.

Overall, the results obtained for common anti-*Pf*CSP antibodies utilizing *IGLV1-47* and different *IGHV* genes demonstrate a common requirement of the Igλ sequence to maintain the overall binding mode. However, the exact antigen binding mode and the effect of single amino acid replacements depend on the sequence of the heavy chain, indicating differences in single residue interactions.

8.4.2 Germline recombination of common anti-*Pf*CSP antibodies utilizing *IGHV3-33*

Antibodies belonging to another important set of public antibodies utilize *IGHV3-33* and *IGKV1-5* genes with a KCDR3 length of 8 amino acids (VH3-33/Vκ1-5/KCDR3:8). They have a considerable affinity to the repeat antigen, as measured by surface plasmon resonance (SPR) although low levels of somatic hypermutations (Figure 16A). The following results highlight the importance of germline recombination events as well as polymorphisms in the *IGHV* locus for the generation of these common anti-*Pf*CSP antibodies.

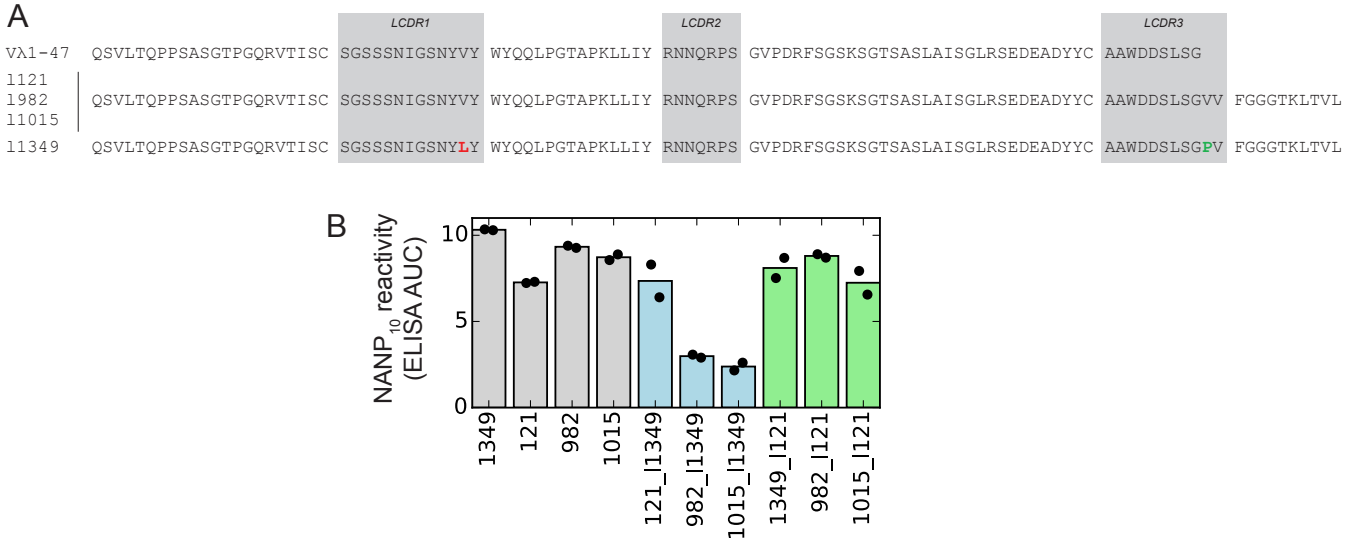


Figure 15. Effect of affinity maturation and light chain recombination on antibody binding is dependent on the heavy chain sequence. A: Sequence alignments of the light chains of antibodies 1349, 121, 982 and 1015. Amino acid exchange due to somatic hypermutation is shown in red. Green: Sequence difference due to alternative VJ recombination. B: CSP reactivity of antibodies 1349, 121, 982 and 1015 as well as chain swap chimeras utilizing 1349l chain and 121l chains.

IGKV gene usage and recombination are essential for antigen binding

Several experiments with chimeric antibodies were conducted to investigate the role of the Igκ chain for antigen binding of VH3-33/Vκ1-5/KCDR3:8 antibodies. When replacing the light chain of representative VH3-33/Vκ1-5/KCDR3:8 antibodies 1210 and 2219 by light chains of either same Vκ segment but longer KCDR3 region the binding was abolished (Figure 16B). Similarly, when replacing the kappa chain of 2219 with a kappa chain with same CDR3 length but utilizing the *IGKV2-28* gene the binding was abolished (Figure 16B).

Moreover when creating chimeric antibodies by swapping Igκ chains within the cluster of VH3-33/Vκ1-5/KCDR3:8, it was found that part of the antigen binding is encoded in the KCDR3 sequence. When replacing the affinity matured Igκ chain of antibody 2140 with other mutated (1210k) or unmutated (2243k) Igκ chains with different recombination sequence (Figure 16D), the

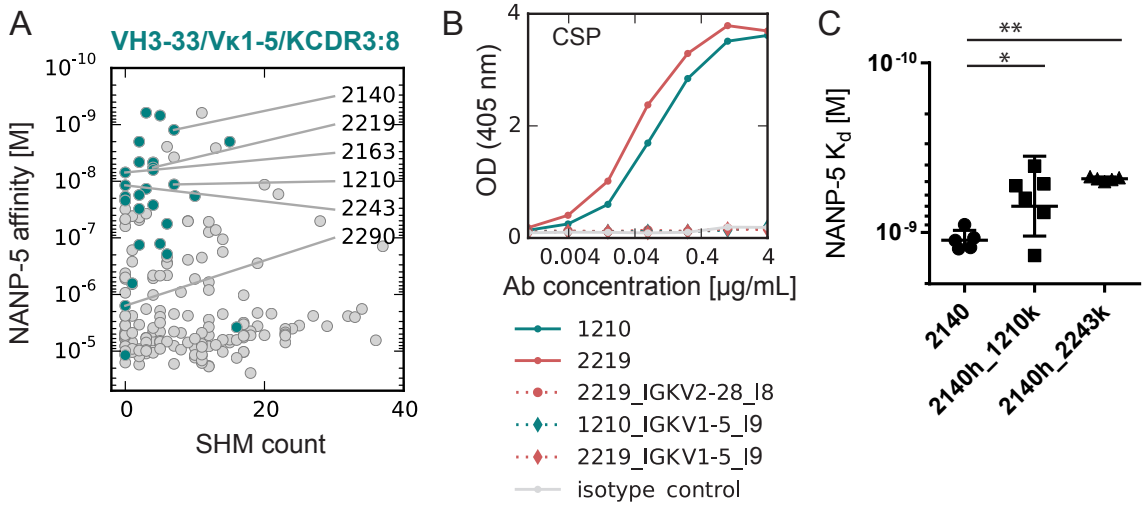


Figure 16. VH3-33/Vκ1-5/KCDR3:8 are high-affine germline-like antibodies and their *Pf*CSP binding depends on the Kappa CDR3 length and sequence. A: $NANP_5$ affinity measured by SPR versus SHM counts for anti-*Pf*CSP antibodies reported in [69]. Antibodies encoded by IGHV3-33 and IGKV1-5 with 8aa-long CDR3 (VH3-33/Vκ1-5/KCDR3:8) are highlighted in green. Representative antibodies are labelled. B: CSP ELISA reactivity of VH3-33/Vκ1-5/KCDR3:8 antibodies 1210 and 2219 with original Igκ chain (solid lines) or after Igκ replacement (dashed lines) encoded by IGKV2-28 with 8aa-long KCDR3 (IGKV2-28_l8) or IGKV1-5 with 9aa-long KCDR3 (IGKV1-5_l9). C: $NANP_5$ affinity measured by SPR of 2140 antibody and 2140h_1210k and 2140h_2243k chimeras. * $p < 0.05$, ** $p < 0.01$, one sided Mann-Whitney test. D: Sequences of 2140, 2243 and 1210 Igκ chains compared to Vκ1-5 germline. The KCDR3 region is shown in pink. Mutated residues are bold and underlined.

affinity could be increased significantly (Figure 16C). This data also demonstrates that in order to generate high antigen binding affinities, kappa chain germline recombination can be more important than affinity maturation, as shown by the higher affinity of 2140h_2243k compared to 2140.

A specific residue encoded in *IGHV3-33* is essential for the generation of these antibodies

*IGHV3-33*01*, the allele utilized in all antibodies of the VH3-33/Vκ1-5/KCDR3:8, strongly resembles other *IGHV3* genes, differing by only one or two amino acid in HCDR2 at positions 50 and 52 (Figure 17A). Residues W52 is exchanged to S or R in *IGHV3-30-3* and some alleles of *IGHV3-33*, *IGHV3-30* and *IGHV3-30-5*. Residue V50 is replaced by F in some alleles of *IGHV3-30-5* and *IGHV3-30*. Using site directed mutagenesis the residues at positions 50 and 52 in antibodies 2140, 1210, 2163, 2219 and 2290 were replaced by the other variants. All tested mutants lost binding to CSP in ELISA (Figure 17B) as well as inhibition potential in *Pf* traversal experiments (Figure 17C).

The presence of the correct allele of *IGHV3-33* in the *IGHV* locus thus seems to be essential for the generation of VH3-33/Vκ1-5/KCDR3:8 antibodies. Since it has been reported that some

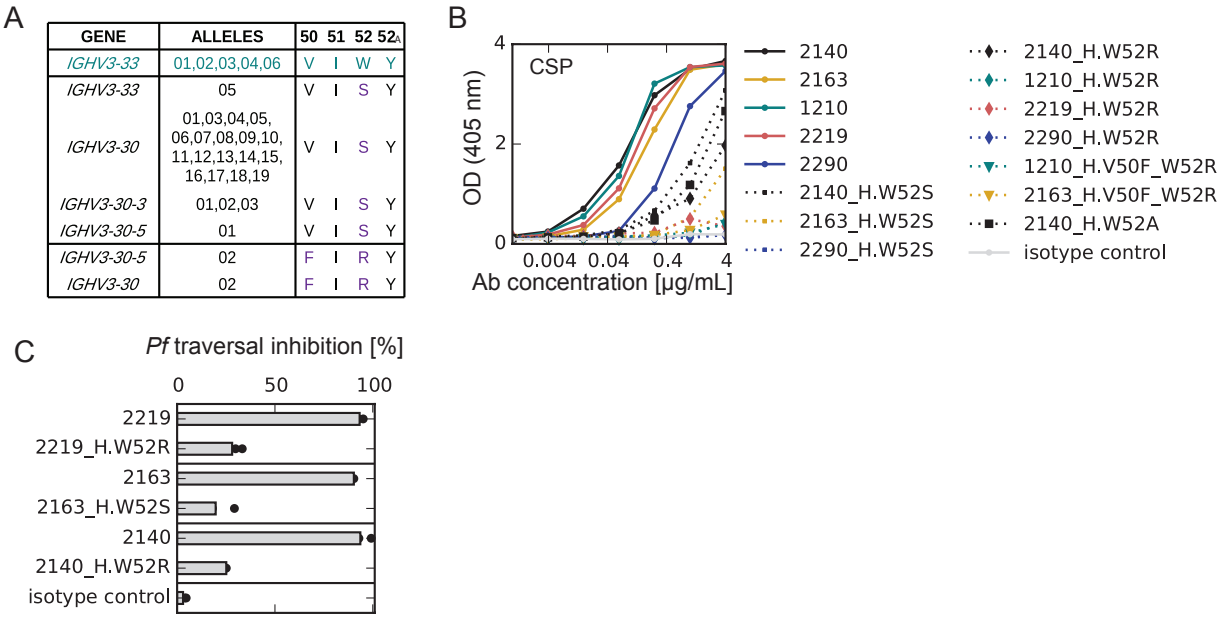


Figure 17. Dependence of VH3-33/Vκ1-5/CDR3 antibodies on specific *IGHV3-33* gene. A: HCDR2 residues 50-52A encoded by different *IGHV3-33*, *IGHV3-30*, *IGHV3-30-3* and *IGHV3-30-5* alleles. Green: *IGHV3-33* allele of VH3-33/Vκ1-5/CDR3 anti-*Pf*CSP antibodies. Purple: Polymorphic residues. B: CSP ELISA reactivity of the indicated antibodies (solid lines) and respective mutants (dashed lines) at position H.V50 and/or H.W52. C: Mean *Pf* liver cell traversal inhibition for the indicated antibodies and respective H.W52 mutants compared to a non *Pf*CSP-reactive isotype control. Dots represent independent experiments.

human haplotypes lack this gene, genomic sequencing of all CHMI donors was used to investigate their genetic predisposition for common anti-*Pf* CSP antibodies. The frequencies of *IGHV3-33** like germline *IGHV* segments in TuCHMI and LaCHMI donors are represented in Figure 18. Within the TuCHMI cohort, one of the patients that failed eliciting VH3-33/Vκ1-5/KCDR3:8 antibodies was also found to lack the germline gene for *IGHV3-33* (Figure 18A, T2_051). However one TuCHMI donor and all three LaCHMI donors for which repertoire sequencing was performed, had *IGHV3-33* in their germline genes but failed to mount the VH3-33/Vκ1-5/KCDR3:8 antibody response (Figure 18A and B), indicating that the presence of a certain germline gene is not sufficient to induce the corresponding antibody response.

Overall the data demonstrates that not only length and sequence of the Igκ chain recombination are crucial for binding of VH3-33/Vκ1-5/KCDR3:8, but also utilization of a very particular variable gene segment of the *IGHV3* family.

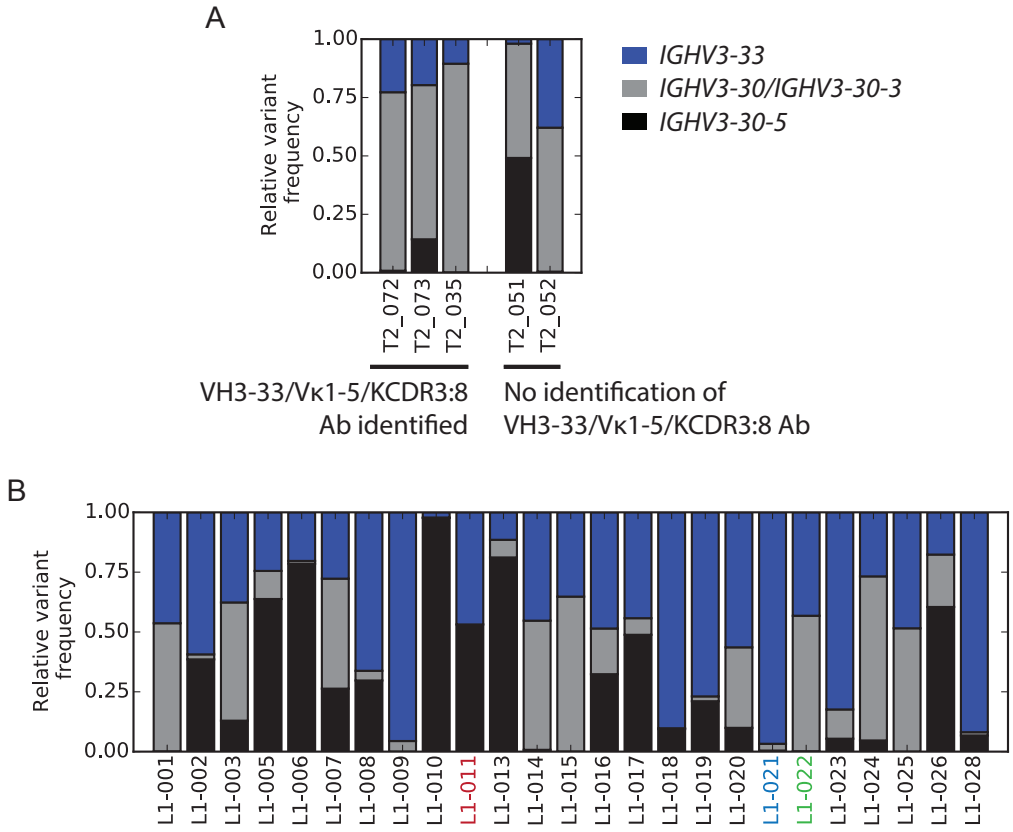


Figure 18. Genomic sequencing of *IGHV3-33-like germline genes in CHMI donors.** The bars indicate the proportion of *IGHV3-33* (blue), *IGHV3-30/IGHV3-30-3* (grey) and *IGHV3-30-5* (black) sequences within the *IGHV3-33**-like sequences (classification according to Figure 17A). A: Sequencing of *IGHV3-33**-like genes from DNA extracted from whole blood for analysed TuCHMI donors. VH3-33/Vκ1-5/KCDR3:8 anti *Pf*CSP antibodies were only identified in 5 out of 8 donors. B: Sequencing of *IGHV3-33**-like genes from DNA extracted from sorted mature naive B cells of LaCHMI donors.

Heterogeneity in the CDR conformation of VH3-33/Vκ1-5/CDR3 antibodies

Binding in the cluster of VH3-33/Vκ1-5/CDR3 is thus strongly constrained by sequence features, that are partly dependent on genetic predisposition. Interestingly the HCDR3, whose sequence is usually very relevant for antigen binding, did not show any obvious motifs or sequence signatures raising the questions of structural conformation of the HCDR3 loop in the set of antibodies.

Molecular homology modelling using Rosetta antibody was used to investigate the conformational landscape of the antibodies. In contrary to what was previously observed for the public *IGLV1-47* antibodies, there was no clear convergence of CDR conformations in the VH3-33/Vκ1-5/CDR3 antibodies, even when using broader clustering parameters (Figure 19). The size distribution of structure similarity clusters is broad and there is no cluster with a contribution of all antibodies, suggesting a rather broad variety of different CDR conformations within the set of VH3-33/Vκ1-5/CDR3 antibodies. This observation is indeed also supported by crystallization experiments of a representative VH3-33/Vκ1-5/CDR3 antibody, 1210, with a NANP-repeat epitope (data not shown, crystal structure kindly provided by J.-P. Julien). In the contact between antibody and antigen, the HCDR3 region is only involved by Cα mediated interactions. Conformational differences in the HCDR3 region could thus be less relevant for antigen binding, as long as the framework provided by KCDR3, HCDR1 and HCDR2 mediates binding.

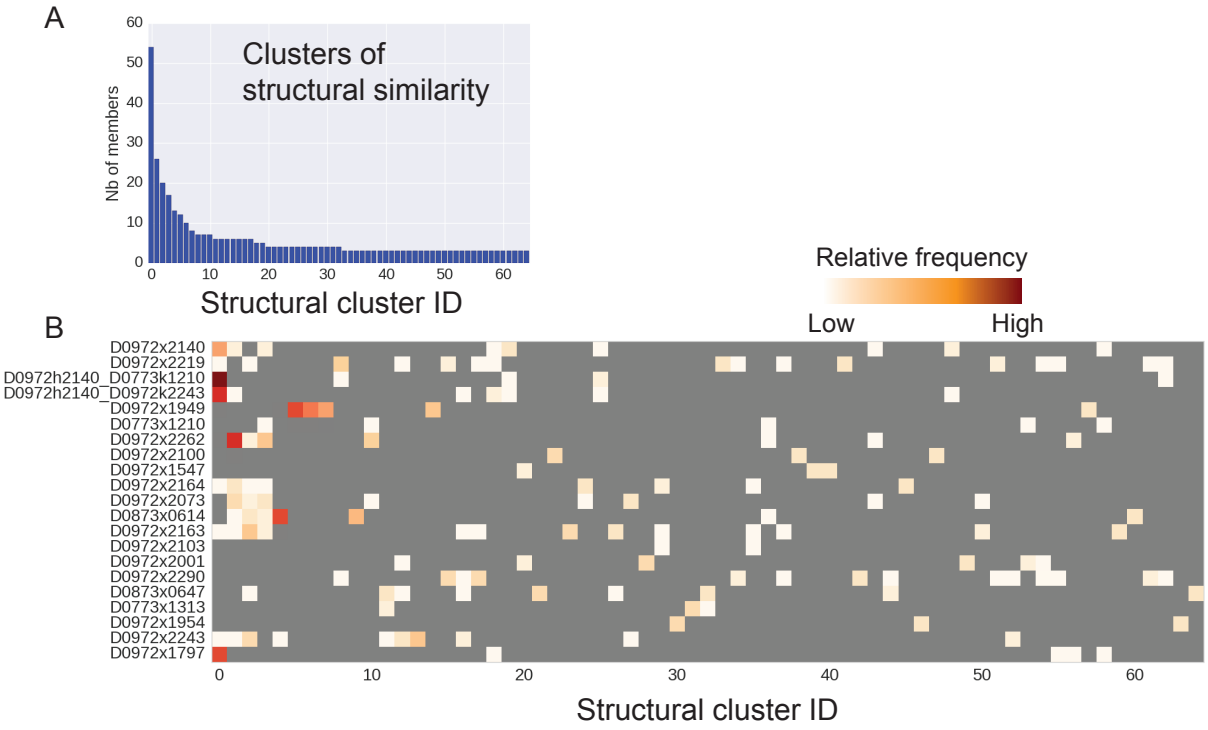


Figure 19. Structural clustering results for VH3-33/Vκ1-5/KCDR3:8 antibodies. A: Size distribution of clusters obtained by GOSSIP clustering. Parameters for GOSSIP clustering: alignment length threshold 0.95, maximal distance between two matching Cα 3.0 Å. B: Contribution of models from different antibodies to clusters of structural similarity.

8.4.3 Identification of a particular binding mode in *IGHV3-33* utilizing antibodies mediated by the tandem-repeat organisation of *Pf*CSP

Affinity maturation is the process that optimizes binding of a B cell receptor to an antigen in a specific binding conformation. If different antibodies share the same binding mode, they should also share similar patterns of affinity maturation. Strong selection patterns were identified when aligning all sequences of the VH3-33/Vκ1-5/KCDR3:8 antibody cluster. Replacement mutations were over-represented in the IgH V gene at positions 31, 50 and 56 as well as in the Igκ V gene at position 93 (Figure 20A). In order to quantify the selection acting at these positions, a computational hypermutation model was used to generate a set of sequences with similar mutation loads as the observed VH3-33/Vκ1-5/KCDR3:8 set [37, 125]. When comparing the amino acid usage of the *Pf*CSP binders (obs) to the reference set (base), a selection of the following replacements became apparent: H.S31N, H.V50I/L, H.N56K/H, K.N93S/N (Figure 20B).

In order to assess the effect of these mutations on affinity maturation, mutants and reversions of two representative antibodies, 1210 and 2163, were generated (section 13.2, appendix). The mutant antibodies were recombinantly expressed and their affinity towards the $NANP_3$ antigen measured in surface plasmon resonance. Exchanges H.S31N, H.V50I were associated with affinity increase against the core epitope in both 1210 and 2163 antibodies (Figure 21). Indeed, the direct involvement of residues at positions H.31 and H.50 was also confirmed in the 1210-$NANP_5$ co-crystal structure (data not shown, crystal structure kindly provided by J.-P. Julien).

However, mutations H.N56K or K.S93N did not lead to significant changes in affinity (Figure 21). Exchanges at positions H.N56 and K.S93 are not selected due to an improved antigen binding by direct interactions with the NANP-repeat peptide. Crystallization of 1210 bound to $NANP_5$ revealed that residues H.K56 or K.N93 are involved in direct interactions of the paratopes of two antibodies

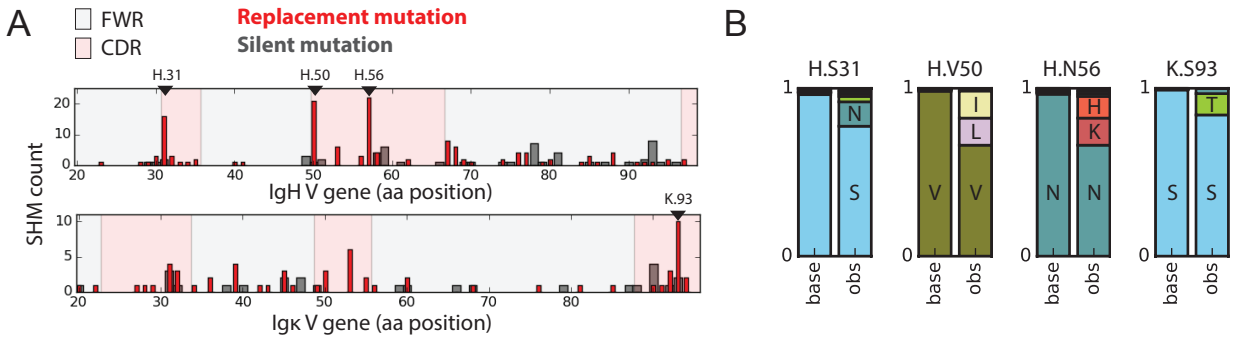


Figure 20. VH3-33/Vκ1-5/KCDR3:8 affinity maturation. A: SHM counts in IgH (top) and Igκ (bottom) chains (n=63). Bars represent the number of silent (grey) and replacement (red) mutations in shaded FWRs (grey) and CDRs (pink) defined according to Kabat scheme. Positions of frequent amino acid exchanges are indicated. B: Observed (obs) aa usage at positions H.31, H.50, H.56 and K.93 compared to a computational baseline model of SHM (base; [37, 125]).

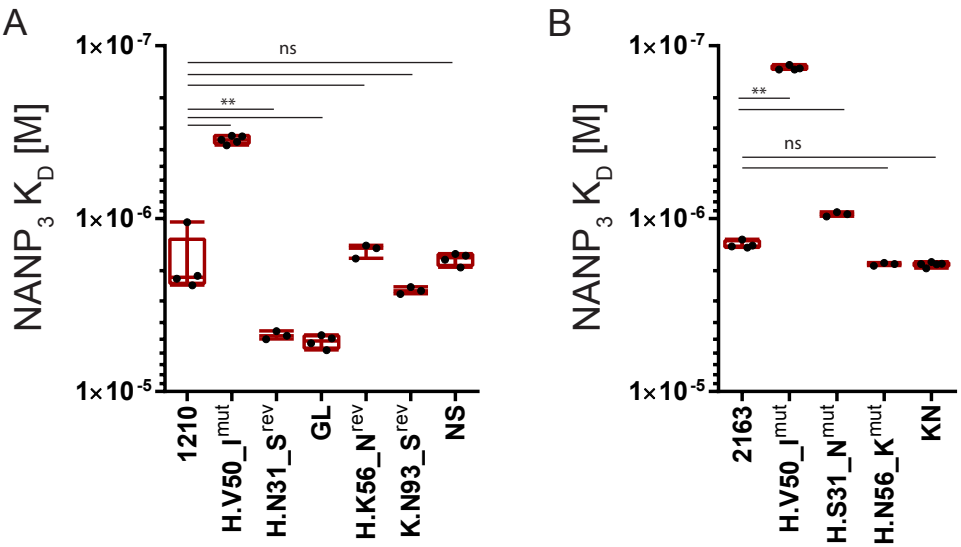


Figure 21. VH3-33/Vκ1-5/KCDR3:8 affinity maturation only partially increases affinity to NANP-repeat. A, B: $NANP_3$ affinity measured by SPR for antibodies 1210 (A) and 2163 (B) and the indicated mutants (Section 13.2, Appendix). Dots represent independent experiments. Multiple testing with Bonferroni correction: ** p $<$0.01, ns: not significant.

binding to neighbouring epitopes in the $NANP_5$ peptide (data not shown, crystal structure kindly provided by J.-P. Julien). This binding mode can be described as paratope-antigen-paratope binding (or sandwich binding) and involves affinity maturation that increases the affinity of a complex antibody-antigen -antibody assembly.

Overall the results on affinity maturation confirm the hypothesis of a common binding mode within the complete set of VH3-33/Vκ1-5/KCDR3:8 antibodies, since similar selection patterns can be observed in all antibodies. Moreover, anti-idiotope affinity maturation is observed due to a common paratope-antigen-paratope binding mode, that is enabled by the repetitive structure of the NANP-repeat.

8.4.4 Antibodies encoded by *IGHV3-23* have the same sandwich binding mode as *IGHV3-33* encoded antibodies

Affinity maturation in VH3-23/Vκ1-5/KCDR3:8 antibodies at positions H.56 and K.93

Anti-idiotope affinity maturation patterns similar to the ones described in VH3-33/Vκ1-5/KCDR3:8 antibodies could also be observed in a set of public antibodies defined by the utilization of *IGHV3-23* instead of *IGHV3-33*. Despite the differences between the two *IGHV* genes, it appears that the binding mode of both types of antibodies is somewhat similar. The antibodies from the *IGHV3-23* cluster have a higher mutation load and lower repeat affinity (Figure 22C) but nevertheless there is a strong selection for amino acid exchanges at positions H.56 and K.93 (Figure 22A). When

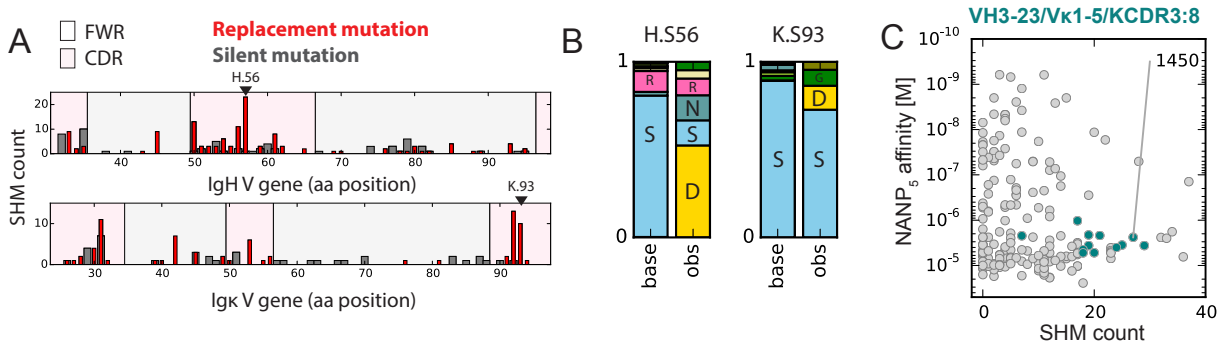


Figure 22. Affinity maturation in common VH3-23/Vκ1-5/KCDR3:8 antibodies. A: SHM counts in IgH (top) and Igκ (bottom) chains (n=24). Bars represent the number of silent (grey) and replacement (red) mutations in shaded FWRs (grey) and CDRs (pink) defined according to Kabat scheme. Positions of frequent amino acid exchanges are indicated. B: Observed (obs) aa usage at positions H.56 and K.93 compared to a computational baseline model of SHM (base; [37, 125]). C: $NANP_5$ affinity measured by SPR versus SHM counts for anti-PfCSP antibodies reported in [69]. Antibodies encoded by IGHV3-23 and IGKV1-5 with 8aa-long CDR3 (VH3-23/Vκ1-5/KCDR3:8) are highlighted in green. Representative antibody 1450 is labelled.

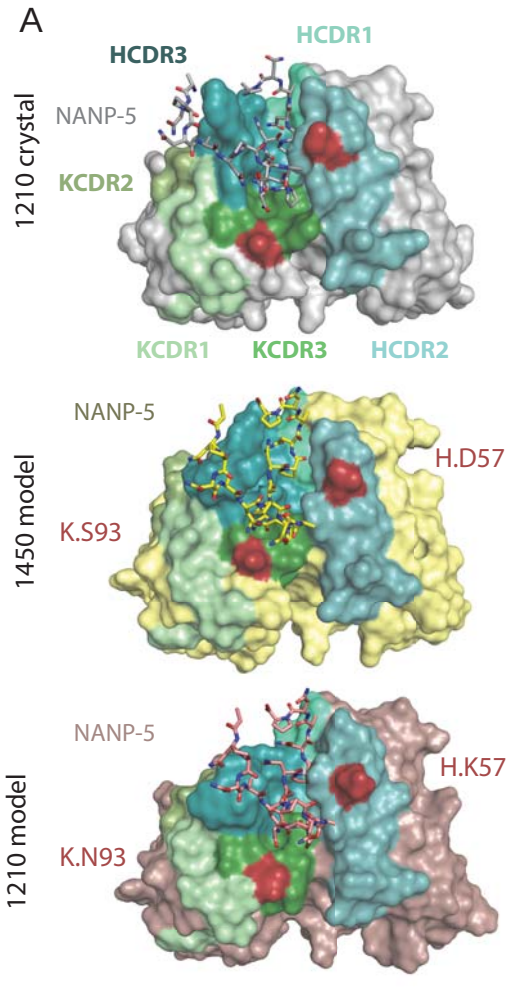
analysing the amino acid usage at these positions, residues with longer polar side chains appear to be selected at both positions. This is similar to what was observed in the *IGHV3-33* utilizing antibodies (Figure 22B).

Binding mode confirmed by molecular modelling

In order to investigate whether these similar affinity maturation patterns were also due to a similar binding mode, the conformation of antigen binding to VH3-23/Vκ1-5/KCDR3:8 antibodies was assessed in order to compare it to the one known for VH3-33/Vκ1-5/KCDR3:8 antibody sandwich binding. Molecular modelling followed by computational docking of the $NANP_5$ repeat was performed for the representative VH3-23/Vκ1-5/KCDR3:8 antibody 1450. As a control for the validity of the method, the same was also done with the representative VH3-33/Vκ1-5/KCDR3:8 antibody 1210 that had been crystallized before in complex with the antigen (structure kindly provided by J.-P. Julien). Figure 23A shows a comparison of the best resulting homology models for 1210 and 1450 antigen binding compared to the crystal structure of 1210-$NANP_5$. The overall antigen binding mode of 1210 and 1450 is very similar.

Figure 23B compares the antibody residues that interact with the antigen in the crystal structure as well as the two docking models. The major differences in residues binding the antigen are observed in the HCDR2 and HCDR3 which are also the regions with highest levels of sequences dissimilarity. A few residues highlighted in grey are conserved between the 1210 and 1450 antigen binding. The importance of residue H.Y58 in 1450 binding was confirmed by site directed mutagenesis and loss of binding of the antibody carrying H.A58 (Figure 23C). The data indicates

that VH3-23/Vκ1-5/KCDR3:8 and VH3-33/Vκ1-5/KCDR3:8 antibodies share similar overall binding mode, but nevertheless differ in the exact set of antigen-antibody interactions. A sandwich binding mode of VH3-23/Vκ1-5/KCDR3:8 antibodies thus seems to be very likely.



B

Region	Position	1210 crystal	1210 model	1450 model
HCDR1	30			S
	31	N	N	N
	32	Y		
	33	G	G	G
	35	H	H	
HFWR2	47		W	
HCDR2	50	V	V	
	52	W	W	S
	52A	Y	Y	G
	53			S
	54			S
HCDR2	56	K	K	
HFWR3	58	Y	Y	Y
HCDR3	95	V	V	E
	96	R		G
	97	D	D	G
	98	S		F
	99	S	S	C
	100	D	D	S
	100A	Y	Y	S
	100B	Y	Y	
	100C	G	G	
	100D			C
	100F			Y
K CDR3	91	Y	Y	Y
	92	N	N	G
	93	N	N	S
	94	Y	Y	S
	95	W	W	W

C

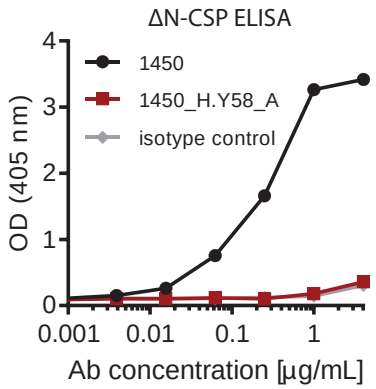


Figure 23. Structure model of a representative VH3-23/Vκ1-5/KCDR3:8 antibody suggest binding mode similar to VH3-33/Vκ1-5/KCDR3:8 antibodies. A: Computational models of antibodies 1450 and 1210 bound to $NANP_5$ compared to the crystalstructure of 1210 bound to $NANP_5$. KCDR1, KCDR2, KCDR3 are shown in light green, green and dark green, respectively. HCDR1, HCDR2 and HCDR3 are colored in light teal, teal and dark teal, respectively. Residues at positions H.56 and K.93 are shown in red. $NANP_5$ residues are shown as stick representation. B: Residues involved in antigen binding in the 1210-$NANP_5$ crystal structure, and in the 1210-$NANP_5$ and 1450-$NANP_5$ models determined using PDBePISA [53]. C: *Pf*CSP ELISA reactivity of antibody 1450 and alanine mutant 1450_H.Y58_A.

9 Discussion

9.1 SciReptor as a tool for annotation of adaptive immune receptor repertoires on single-cell level

An essential step in this work was to adapt the existing Ig sequencing protocol to the Illumina sequencing technology and to develop a bioinformatics pipeline that would make it easy to access the sequence annotation. The sciReptor software was developed and successfully used to analyse the raw sequencing data for all B cell repertoires presented in this work. The efficiency of successful computational sequence annotation was comparable to the efficiencies estimated by Sanger sequencing [68]. After its first publication sciReptor has further been developed to analyse T cell repertoires (Bartsch and Busse, unpublished work).

Recent developments in the field of immune receptor sequencing have highlighted the importance of correct genomic databases for accurate annotation of germline genes and somatic hypermutations. Several new computational tools allow de novo allele prediction based on repertoire data sets [22, 37]. However, in the case of the here studied datasets, sequence and cell numbers are not high enough to reliably reconstruct germline gene sequences. If germline gene reconstruction prior to repertoire annotation was to become relevant, bulk sequencing could be applied to generate a database of germline genes which could then be plugged in into the library module of sciReptor.

In contrary to the widely used bulk sequencing approaches, sequencing of B cell repertoires at single-cell level allows correct assessment of clonality as well as identification of paired heavy and light chain genes. Especially the latter is essential when exploring public antibody repertoires. In the field of adaptive immune receptor repertoire (AIRR) sequencing sciReptor remains one of the few tool designed to analyse repertoires at single-cell level, and capable of integrating non-sequence annotation like cellular phenotypes for example. Other common AIRR analysis tools have integrated consensus building functions [37, 107], but none of them allows full integration of annotations at single-cell level. An additional major challenge in the assessment of antibody repertoires remains the lack of high-throughput expression systems that would allow systematic testing of antibody reactivity. The here presented work highlights the necessity of testing antibody reactivities, especially when investigating public repertoires. First attempts on high-throughput affinity measurements have been made [2], but they are by far not commonly usable.

9.2 Pre-erythrocytic immune response response in LaCHMI donors

An important heterogeneity of pre-erythrocytic immune responses was observed in the Gabonese donors during the course of the LaCHMI study. Not only was the study outcome in terms of disease and symptom development very different from donor to donor, but also the humoral immune parameters differed a lot.

9.2.1 Heterogeneity of the study outcome is due to heterogeneity of pre-erythrocytic immunity

Out of the 8 Gabonese LaCHMI subjects that were completely protected against the infection, only 3 showed relevant anti-*Pf*CSP titers before the start of the study on day C-1. The malaria history of the Gabonese subjects recruited into the LaCHMI trial is difficult to judge and thus it is not clear whether the low frequency of pre-existing pre-erythrocytic immunity and the even lower frequency of *Pf*CSP serum antibodies is due to lack of sufficient previous exposure or due to heterogeneity of the immune responses upon repeated exposure. However low frequencies of anti-*Pf*CSP responses in naturally exposed individuals have been described in the endemic areas, including the region the LaCHMI study participant were living in [46, 104]. According to a study, that compared rates of seroconversion for different antigens over 10 years, the reduced frequencies of anti-*Pf*CSP serum antibody titers may be due to lower exposure frequencies [116].

Interestingly, the donors that have an active immune response and increase the levels of anti-*Pf*CSP titers in serum from day C-1 to day D28 are the ones that are also treated early on in the infection. It has been reported that the development of blood stage parasites and acute malaria is detrimental for the successful development of pre-erythrocytic humoral immunity [50]. This hypothesis could explain the lack of anti-*Pf*CSP titers in donors that develop asymptomatic parasitemia and also present a mechanism of why under natural exposure only few individuals mount anti-*Pf*CSP memory B cell responses. However, this hypothesis does not give sufficient explanation for why the protected donors that eliminate the parasite before blood stage parasitemia do not raise anti-*Pf*CSP titers.

The cases of pre-erythrocytic immunity without presence of anti-*Pf*CSP serum titers confirm results from previous studies in naive donors [70] and illustrate the presence of other non antibody dependent and non *Pf*CSP dependent mechanisms for immunity against sporozoite infection. The presence of a strong plasmablast response without raising levels of anti-*Pf*CSP titers, as observed in donors L1-002 or L1-009, highlights that the humoral immune response can also be directed against other parasite antigens. Non humoral mechanisms of pre-erythrocytic immunity include CD8 T cells or interferon-gamma secreting CD4 T cells that act against the *Pf* liver stages or EEFs [64, 94, 95].

9.2.2 Failure of identification of *Pf*CSP reactive antibodies in LaCHMI participants

Several cloning attempts failed to identify antibodies with measurable *Pf*CSP reactivity in ELISA in the LaCHMI cohort, although the donors L1-011, L1-021 and L1-022 were selected for repertoire analysis based on their humoral anti-*Pf*CSP profile. The results correspond well to the observations of a previous study in *Pf* experienced donors, where the frequencies of *Pf*CSP reactive B cells were also extremely low [104].

The selection of Gabonese donors for repertoire sequencing was based on the assumption that anti-*Pf*CSP titers also correlate with circulating *Pf*CSP specific memory B cells. The failure to isolate *Pf*CSP reactive antibodies from the sorted cells however could be due to the fact that only very few B cells with measurable affinity were circulating in the peripheral blood of the selected donors at the time of blood sampling. The screening approaches then failed to identify them because they were too rare. Indeed, the three donors selected for repertoire analysis did not show an increase of anti-*Pf*CSP serum titers, but rather even tended to decrease them, which suggests that there was no active anti-*Pf*CSP B cell response after sporozoite injection. It is possible that the donors had few reactive memory B cells in the circulating blood and that the serum titers were maintained by plasma cells residing in the bone marrow for example. *Pf*CSP exposure under natural conditions is with much lower antigen quantities than the repeated immunization during the TuCHMI study e.g. and thus the memory responses cannot be compared. Retrospectively, in order to find higher numbers of circulating *Pf*CSP specific B cells it would probably have been better to choose the donors with active ongoing anti-*Pf* CSP B cell response, i.e. with increasing anti-*Pf*CSP titers during the infection.

Even if the previous hypothesis is reasonable, there are several indications for an *Pf*CSP specific population within the B cell memory of the selected donors. First of all, the enrichment of *Pf*CSP reactive B cells in FACS to a level of 2% (Figure 6) compared to a baseline level of approximately 0.02 % in healthy donors [104] indicates the presence of a truly *Pf*CSP specific population. In addition, AtM B cells were found to be enriched in the *Pf*CSP binding B cell population and showed a phenotype similar to the one previously described for AtM. This observation indicates that the *Pf*CSP binding of cells in FACS is not an artefact but that *Pf*CSP-specific B cells are present in the AtM population. This population has indeed previously been described to be reactive to pathogen antigens [23, 66].

The lack of identification of *Pf*CSP specific B cells is thus more likely do to a lack of sensitivity of the reactivity assessment. Antibodies with a very low *Pf*CSP affinity, may still bind *Pf*CSP in FACS, but the interaction may not be detectable any more in the less sensitive ELISA method. Supporting this idea, many of the BCRs that bound *Pf*CSP in FACS had an IgM isotype, a phenomenon that has also been observed in the anti-*Pf*CSP repertoire of European CHMI donors [69], but not in antibody repertoires directed against viral antigens for example [51]. However the antibody reactivity was tested on antibodies cloned and expressed as IgG1 isotypes, which in contrast to the pentameric IgM

only has two antigen binding sites and may thus be less useful to detect antigen binding in which avidity plays a major role. In addition, selection of antibodies with low affinity to *Pf*CSP could be favoured by the repetitive structure of the protein or by anti-idiotope interactions that select BCRs that only have weak affinity to the antigen epitope but higher affinity in the complex assembly. In other words, anti-idiotope interactions could lower the threshold of antibody-antigen binding affinity that is necessary to activate a specific B cell.

In summary, the LaCHMI participants showed very heterogeneous pre-erythrocytic immune responses, especially to the major sporozoite antigen *Pf*CSP. Moreover, under natural exposure conditions, it seems not only to be difficult to mount an anti-*Pf*CSP response at all, but it also seems to be difficult to recruit highly reactive antibodies against *Pf*CSP. The present work suggests that many B cells with low affinity are recruited, which may be due to the repetitive structure of the *Pf*CSP antigen in combination with interactions between antibodies binding to neighbouring epitopes.

9.3 Technical consideration on public antibody repertoires and the impact of sequence similarity

A sequence similarity clustering approach was used to identify the public B cell repertoire after CHMI, i.e. B cells encoding antibodies with similar sequence and supposedly also similar specificity. The approach led to an identification of a set of small and very diverse clusters of public antibodies encoded by very different heavy and light variable gene combinations. Focussing on anti-*Pf*CSP antibodies, the relationships between sequence, structure and antigen binding were further investigated to describe the characteristics of the public anti-*Pf*CSP repertoire.

9.3.1 Predicting antibody binding based on sequence similarity should be confirmed by reactivity testing

The predictive clustering approach presented in Figure 11 demonstrates that sequence similarity usually also correlates with similar reactivity. However, only the reactive antibodies from European donors were successfully clustered using this approach. For the African donors, only few antibodies clustered and none of the antibodies that were cloned also actually bound *Pf*CSP in ELISA. Overall, the clustering approach seems to be sensitive and specific enough when clustering repertoires that contain the desired antibodies at sufficient frequency. For the Gabonese repertoires, the frequency of *Pf*CSP specific B cells was too low and thus too many false positive antibodies were identified. It can be overall concluded that basing the definition of public repertoire only on sequence similarity is difficult and should not be done without confirming the reactivity of the antibodies. Although most of the public repertoire studies rely on sequence data only, recent studies start to include small scale reactivity data, which is still limited by the lack of high-throughput expression systems [44].

9.3.2 The accuracy of public repertoire identification is influenced by the choice of the clustering algorithm

The technical implementation of antibody clustering in this work had specific problems related to the processes which lead to the generation of an immunoglobulin sequence. In the here presented approach the complete antibody sequence starting from CDR1 to the end of the J segment was aligned, giving high importance to the long germline encoded variable gene and less importance to the shorter but highly diverse CDR3 regions generated by recombination. Antibodies with low levels of somatic hypermutations eventually clustered together although they had very little HCDR3 similarity and not the same reactivity. This was probably the case for the antibody with VH3-33/Vκ1-5/KCDR3:8 sequence characteristics that was identified in one of the Gabonese donors but did not bind even when the critical mutations were reverted. Recognition of different antigens that is only based on HCDR3 differences has indeed also been described in detail in a previous study [102]. In addition to specificity problems, the approach is also not well suited to identify antibodies with similar antigen binding sites but high levels of somatic hypermutation, since replacements at positions irrelevant for antigen binding diversify the antibody sequence.

The results obtained using a simple full length amino acid sequence alignment were satisfying, but could however be improved by enhancing the clustering algorithm. Similar binding relies mostly on single residues that are exposed to the antibody surface and even though the antibody framework can be quite different, these residues can still be exposed to the surface in a similar configuration. To take into account this knowledge, it would be useful to combine structure modelling and alignment with sequence alignment.

9.3.3 Sequence similarity only partly entails structure similarity

In addition to reactivity assessment of public anti-*Pf*CSP antibodies, structure modelling was used to assess the structural diversity within two sets of antibodies with high sequence similarity. The HCDR3 is the antibody loop with the highest sequence and structure variation and it is thus also the one that is most extensively modified during the homology modelling in the Rosetta software package. Models with high similarity to the actual protein structure are usually identified by structural convergence within the models with the best scores. In this work, antibodies utilizing *IGLV1-47* showed convergent structure clustering, whereas VH3-33/Vκ1-5/KCDR3:8 antibody models did not indicate any structural convergence.

It can be concluded, that HCDR3 in the *IGLV1-47* utilizing antibodies have a similar conformations which might also be due to a shorter length and due to the presence of a proline residue within the HCDR3 region. Shorter HCDR3 loops are less flexible and proline residues are known to rigidify protein loops. The HCDR3 of VH3-33/Vκ1-5/KCDR3:8 antibodies might simply be less rigid and play a minor role in the interactions with the NANP-repeat. The minor contribution of HCDR3

in antigen binding of VH3-33/Vκ1-5/KCDR3:8 was also confirmed by crystallization experiments conducted by J.-P. Julien. The current literature concerning the structure evolution of antibodies during the course of affinity maturation reports structure convergence that goes along with sequence convergence [92]. However affinity maturation does not typically result in rigidification of antibody structures [49], which supports the idea of high affinity being mediated even with flexible HCDR3 regions.

The literature as well as the here presented data on molecular characterization of binding interactions indicates many different ways in how binding can be accomplished, ranging from isolated binding motifs to complex interaction networks of antibody and antigen. As discussed above, it is difficult to find an adequate clustering algorithm for public antibodies, and resulting hypothesis should be confirmed by reactivity testing. Sequence similarity does not always reflect structure similarity, and, most importantly, similar binding can be achieved even with small differences in structure, as demonstrated by the data on VH3-33/Vκ1-5/KCDR3:8 antibodies.

9.4 Characteristics of the common anti-*Pf*CSP B cell repertoire after CHMI

Based on the observations from the predictive cloning of anti-*Pf*CSP specific antibodies, the clusters of public antibodies identified in the European CHMI donors also share similar reactivity. The common anti-*Pf*sporozoite repertoire is surprisingly heterogeneous and is composed by several small and diverse clusters of similarity. In contrary to that, independent studies on B cell repertoires after influenza vaccinations have reported stronger convergence within antibody rearrangements and more homogeneous antibody response between donors [48, 110]. One explanation for this phenomenon is certainly that immunization with a complete parasite confronts the immune system with a diverse set of antigenic stimuli and that B cells with many different specificities are activated.

However, also the subset of the public repertoire specific to *Pf*CSP seems to be rather diverse. The induction of a broad and diverse response of antibodies could be due to the high abundance of antigen during the immunization combined with a low frequency of *Pf*CSP specific mature naive B cells. The public antibodies with high affinity to the NANP-repeat have sequence characteristics that likely make them rare in generation of the repertoire (KCDR3 length of 8 amino acids, *IGHV* gene restriction e.g.). Antibody targets in other diseases may be easier to bind or be able to activate a higher number of corresponding specific mature naive B cells. The influence of mature naive B cell precursor frequency on the development of the memory B cell repertoire to a specific antigen has recently been illustrated in a mouse system with a VRCO1-class germline targeting HIV antigen [1].

9.4.1 *IGHV3* prevalence within public anti-*Pf*CSP antibodies

It appears that clusters of common anti-*Pf*CSP antibodies frequently utilize IgH genes from the *IGHV3* family, a claim that was already made when looking at the complete anti-*Pf*CSP repertoire [69]. In the same lines, a study on RTS,S induced antibodies [75] as well as a study on naturally induced anti-*Pf*CSP antibodies [104] also identified antibodies utilizing *IGHV3* genes. A preferred usage of *IGHV3* heavy variable genes was equally described for bacterial polysaccharides [3], whereas hemagglutinin of influenza frequently is bound by antibodies encoded by the IGHV1 [7, 48, 110] and antibodies in celiac disease often utilize IGHV5-51 [88].

The *IGHV* family preferentially found in antibodies against a certain target is most likely influenced by the structure complementarity of the resulting antibody with the antigen. The resolved crystal structures of antibodies binding to NANP-repeat peptides show binding of the peptide in a concave pocket, that is formed by several CDRs of the antibody ([75, 104], crystal structure provided by J.P. Julien). The here presented data on molecular characteristics of public anti-*Pf*CSP antibodies confirms, that several regions of the antibody are important for antigen binding. In contrast to this, in antibodies binding the receptor binding site of influenza hemagglutinin for example single motifs within the HCDR3 interact with the antigen [93, 124]. Several very different germline genes can encode these antibodies. Similar observations were made for antibodies binding the HIV envelope V2 Apex, where antigen binding is mediated by a motif in the HCDR3 [6].

Interestingly, it has been reported that concavity of the paratope as for example determined by CDR lengths influences the target preference between globular and disordered antigens [21, 59]. The different *IGHV* genes are known to have different intrinsic lengths of CDR region, which can adopt different conformations according to their amino acid conformation. A new clustering approach gives good overview over the different loop configurations found in antibody CDRs [71]. According to this clustering, interestingly *IGHV3-33*, *IGHV3-23* and *IGHV4-59* have similar CDR conformations, which are also the genes enriched in the public repertoire (Figure 10).

9.4.2 The public anti-*Pf*CSP repertoire is influenced by genetic polymorphisms

In addition to the specificity and binding, the diversity and similarity of public antibody repertoires can also be determined by the immunoglobulin genes present in the germline of every individual. The here described VH3-33/Vκ1-5/KCDR3:8 antibodies depend on a *IGHV* germline gene that is absent in some individuals. A previous study on complete *IGHV* haplotype sequences reported that *IGHV3-33* is located in a locus of major structural polymorphism within the human *IGHV* locus [114, 115]. Genomic or allelic predisposition to mount a certain type of antibodies has also been discussed for antibodies in different viral infections [7, 58, 127]. Other genetic polymorphisms that lead to differences in B cell repertoire generation could be copy number variations of specific germline genes [7] or polymorphisms in regulatory regions of the immunoglobulin loci.

Nevertheless, genetic predispositions only explain part of the variability that is observed in antibody repertoires. A twin study on acute responses to Varicella-Zoster virus vaccination reported that twins share similarities in the overall memory B cell compartment, but that stochastic effects dominate the selection of clones in acute immune responses [111]. In the case of hepatitis B vaccination, a well-known example for incomplete humoral immune responses in a certain number of patients, it has been hypothesised that HLA restriction might play a role [113] or that the ratio of Th1- and Th2-type T cells plays a role in the generation of the B cell response [25].

In summary, the present work highlights two possible explanations for the observed heterogeneity of the common anti-*Pf*CSP repertoire: low precursor frequencies of *Pf*CSP-specific B cells and genetic polymorphisms. The apparent enrichment of a specific *IGHV* family may be due to the conformation and length of germline encoded CDRs.

9.5 Antigen binding and affinity maturation against a repetitive epitope

The most surprising and interesting result of the present work is the observation that antibodies directed against *Pf*CSP directly interact with each other. The interaction of the two antibody molecules, as determined in the crystal structure by J.P. Julien, is absolutely symmetrical in terms of antigen binding, which means that the interactions between antibody and antigen are identical in both Fabs contributing to the sandwich. Direct interactions of antibodies with themselves at the antigen binding site have previously been observed in two cases. In 1996 an anti-idiotype antibody was described, that forms dimers, but without the simultaneous involvement of antigen binding [8]. In 2016, another histone-specific antibody was reported to bind its antigen in a sandwich conformation, but the antigen interaction with "upper" and "lower" sandwich Fab are non-symmetric [39]. The symmetry of the here described anti-*Pf*CSP antibody binding is only possible due to the repetitive sequence and structure of the NANP-repeat peptide.

9.5.1 Antigen binding conformation leads to anti-idiotope affinity maturation

The selection patterns of anti-idiotope affinity maturation are one of the observations that indicated a common *Pf*CSP binding mode that involves direct interaction of neighbouring antibodies. Ca. 50% of the affinity maturation observed in VH3-33/Vκ1-5/KCDR3:8 antibodies is directed against the antibody paratope of a neighbouring antibody and thus improves the assembly of an antigen with several antibodies. In VH3-23/Vκ1-5/KCDR3:8 antibodies, this percentage might even be higher since no other hotspots of affinity maturation become apparent in the sequence alignments. This type of self-dependent affinity maturation of BCRs has not been described before. Although there exist reports on autonomous signalling in CLL B cells, where BCRs directly interact with each other and affinity mature against them-self [61], this phenomenon has never been associated with improved binding to antigens.

9.5.2 Potential biological mechanisms for anti-idiotope affinity maturation

Mechanistically the anti-idiotope affinity maturation could rely on two different processes: antibody feedback of antibodies binding to *Pf*CSP antigen in germinal centres or direct improvement BCRs cross-linking on the surface of a B cell. Antibody feedback has been primarily reported to mask epitopes on antigens and thus restrict the set of available BCR targets in germinal center reactions [99, 132]. However it is also possible that antibodies binding to antigen in germinal centers might extend the epitopes targeted by newly generated B cells and antibodies, as in anti-idiotope affinity maturation. Anti-idiotope affinity maturation due to direct improvement of BCR cross-linking on the B cell surface is also conceivable, although it would have to be clearly shown that this is sterically possible. In order to investigate the two hypothesis, experiments could be done to measure the capacity of different antibody mutants (with or without anti-idiotope mutations) to activate B cells in vitro, as e.g. calcium flux experiments.

While the antigen binding of the two Fabs presented in the crystal structure is symmetric, the interaction between the two antibody variable regions is asymmetric (data not shown, crytal structure kindly provided by J.-P. Julien). Especially in the case of interaction between BCRs from the same cell, this aspect could induce competition of different mutations and lead to inefficient affinity maturation over a longer period of time. In order to further investigate this, experiments in genetically modified mice, that only express one BCR of VH3-33/Vκ1-5/KCDR3:8-type, could be performed. The evolution of the BCR sequences could be monitored throughout multiple immunizations with *Pf*CSP.

9.5.3 Role of the repetitive *Pf*CSP structure in antibody binding

The NANP-repeat structure of *Pf*CSP is essential to enable the direct interaction of two antigen binding sites. Especially the type-I and type-II β turns that are facilitated by the proline residues and that alternate to form a loop structure are essential to bring two antibody molecules into close proximity. None of the other *Plasmodium* species has the exact same repetitive pattern, but the CSP protein of *Plasmodium berghei* e.g. uses several repetitive motifs that include proline but not the N and A amino acids that form stabilizing hydrogen bonds in the *Pf*CSP repeat. It is thus unlikely that antibodies directed against the CSP of other *Plasmodium* species form the same type of complex. Nevertheless also binding to the repetitive molecule at other angles could induce direct interactions of antibodies. Similar to that, other repetitive antigens like polysaccharides or viral capsule proteins could induce non-antigen directed affinity maturation by allowing interactions between antibodies or BCRs binding neighbouring epitopes.

In summary, the observations on anti-idiotope interaction and affinity maturation of anti*Pf*CSP antibodies highlight the importance of the repetitive structure of the NANP-repeat in immune responses against the parasite. It is important to note, that the here presented results broaden the

understanding of antibody binding to *Pf* CSP. It was previously postulated that binding of antibodies to *Pf* CSP might have considerable entropic cost due to the flexible structure of *Pf* CSP and necessary stabilization of the antigen [26]. However, the principle of anti-idiotope affinity maturation shows that the complex assembly of several antibodies to the antigen might actually be beneficial and increase antibody avidity by cooperative effects.

10 Outlook

The discovery of anti-idiotope binding and affinity maturation of antibodies binding to the repeat domain of *Pf*CSP raises several novel questions that will have to be addressed in future research. First, it will be important for the immunological community to investigate whether this type of complex assembly also plays a role in other repetitive antigens like bacterial polysaccharides. In this context it will be necessary to understand whether anti-idiotope binding can induce or facilitate T cell independent responses.

Second, it will be crucial to delimit the exact effect of such interactions on in vivo immune reactions in order to find out whether the interaction is used by the parasite as a kind of immune evasion mechanism or whether it simply makes the B cell response to the antigen more efficient. The parasite could for example use the repetitive region of *Pf*CSP to induce anti-idiotope interactions in antibodies and reduce the affinity threshold of B cells that it activates. To test this, B cell activation experiments with BCRs identified in the Gabonese donors could be performed to test whether these BCRs are at all activated by *Pf*CSP. In line with the theory of parasite immune evasion, preliminary results on parasite inhibition in vitro and in vivo indicate that anti-idiotope binding does not play a role in parasite neutralization. Sterical hindrance due to high *Pf*CSP density on the parasite surface could for example impair efficient anti-idiotope interactions on the parasite surface. All affinity maturation that increases anti-idiotope affinity would then have no effect on parasite inhibition.

With the idea of exploiting anti-idiotope binding to target specific germline recombinations in vaccination, we have recently filed a patent for a chimeric complex of a VH3-33/Vκ1-5/KCDR3:8 Fab covalently linked to a $NANP_5$ repeat. This construct might be used to specifically target the germline precursor B cells that lead to high-affine and protective antibodies against malaria. However, since it was shown in this work, that structural polymorphisms of the *IGHV* locus might influence the generation of anti*Pf*CSP antibodies, this approach has to be investigated thoroughly. In case the chimeric complex could only target VH3-33/Vκ1-5/KCDR3:8 B cells specifically, only such individuals with *IGHV3-33* in their germline DNA would benefit from the vaccination.

Finally, in the same lines, it would be interesting to investigate the population genetics of *IGHV3-33* deletions. Since this gene is used in high-affine germline-like anti-sporozoite antibodies, having the gene could have an advantage for malaria survival. The current human genome databases however do not allow clear statements about *IGHV3-33* frequencies in European and Africans for example, because the annotations are not yet precise enough to differentiate between the different very similar germline genes.

11 References

[1] R. Abbott, J. Lee, S. Menis, P. Skog, M. Rossi, T. Ota, D. Kulp, D. Bhullar, O. Kalyuzhniy, C. Havenar-Daughton, W. Schief, D. Nemazee, and S. Crotty. Precursor Frequency and Affinity Determine B Cell Competitive Fitness in Germinal Centers, Tested with Germline-Targeting HIV Vaccine Immunogens. *Immunity.*, S1074-7613(17)30524-1, 2017. doi:10.1016/j.immuni.2017.11.023.

[2] R. Adams, T. Mora, A. Walczak, and J. Kinney. Measuring the sequence-affinity landscape of antibodies with massively parallel titration curves. *Elife.*, 2016. doi:10.7554/eLife.23156.

[3] E. Adderson, P. Shackelford, A. Quinn, P. Wilson, M. Cunningham, R. Insel, and W. Carroll. Restricted immunoglobulin VH usage and VDJ combinations in the human response to *Haemophilus influenzae* type b capsular polysaccharide. Nucleotide sequences of monospecific anti-*Haemophilus* antibodies and polyspecific antibodies cross-reacting with self antigens. *J Clin Invest*, 91(6):2734–43, 1993. doi:10.1172/JCI116514.

[4] M. Aikawa, N. Yoshida, R. Nussenzweig, and V. Nussenzweig. The protective antigen of malarial sporozoites (*Plasmodium berghei*) is a differentiation antigen. *J Immunol*, 126(6):2494–2495, 1981. URL `http://www.jimmunol.org/content/126/6/2494`.

[5] R. Amino, S. Thiberge, B. Martin, S. Celli, S. Shorte, F. Frischknecht, and R. Ménard. Quantitative imaging of *Plasmodium* transmission from mosquito to mammal. *Nat Med*, 12(2):220–4, 2006. doi:10.1038/nm1350.

[6] R. Andrabi, J. Voss, C. Liang, B. Briney, L. McCoy, C. Wu, C. Wong, P. Poignard, and D. Burton. Identification of Common Features in Prototype Broadly Neutralizing Antibodies to HIV Envelope V2 Apex to Facilitate Vaccine Design. *Immunity*, 43(5):959–973, 2015. doi:10.1016/j.immuni.2015.10.014.

[7] Y. Avnir, C. Watson, J. Glanville, E. Peterson, A. Tallarico, A. Bennett, K. Qin, Y. Fu, C. Huang, J. Beigel, F. Breden, Q. Zhu, and W. Marasco. IGHV1-69 polymorphism modulates anti-influenza antibody repertoires, correlates with IGHV utilization shifts and varies by ethnicity. *Sci Rep*, 6:20842, 2016. doi:10.1038/srep20842.

[8] N. Ban, J. Day, X. Wang, S. Ferrone, and A. McPherson. Crystal structure of an anti-anti-idiotype shows it to be self-complementary. *J Mol Biol.*, 255(4):617–27, 1996.

[9] G. Bastiaens, M. van Meer, A. Scholzen, J. Obiero, M. Vatanshenassan, T. van Grinsven, B. Sim, P. Billingsley, E. James, A. Gunasekera, E. Bijker, G. van Gemert, M. van de Vegte-Bolmer, W. Graumans, C. Hermsen, Q. de Mast, A. van der Ven, H. S, and R. Sauerwein. Safety, Immunogenicity, and Protective Efficacy of Intradermal Immunization with Aseptic, Purified, Cryopreserved *Plasmodium falciparum* Sporozoites in Volunteers Under Chloroquine Prophylaxis: A Randomized Controlled Trial. *Am J Trop Med Hyg*, 94(3):663–73, 2016. doi:10.4269/ajtmh.15-0621.

[10] M. Behet, G. Foquet, Land van Gemert, E. Bijker, P. Meuleman, G. Leroux-Roels, C. Hermsen, A. Scholzen, and R. Sauerwein. Sporozoite immunization of human volunteers under chemoprophylaxis induces functional antibodies against pre-erythrocytic stages of *Plasmodium falciparum*. *Malar J*, 13:136, 2014. doi:10.1186/1475-2875-13-136.

[11] J. Beier, J. Davis, J. Vaughan, B. Noden, and M. Beier. Quantitation of *Plasmodium falciparum* sporozoites transmitted in vitro by experimentally infected *Anopheles gambiae* and *Anopheles stephensi*. *Am J Trop Med Hyg*, 44(5):564–70, 1991. doi:10.4269/ajtmh.1991.44.564.

[12] C. Berek and C. Milstein. Mutation drift and repertoire shift in the maturation of the immune response. *Immunol Rev*, 96(1):23–41, 1987. doi:10.1111/j.1600-065X.1987.tb00507.x.

[13] G. Bijker, EMand Bastiaens, A. Teirlinck, G. van Gemert, W. Graumans, M. van de Vegte-Bolmer, R. Siebelink-Stoter, T. Arens, K. Teelen, W. Nahrendorf, E. Remarque, W. Roeffen, A. Jansens, D. Zimmerman, M. Vos, B. van Schaijk, J. Wiersma, A. van der Ven, L. de Mast, Qand van Lieshout, J. Verweij, C. Hermsen, A. Scholzen, and R. Sauerwein. Protection against malaria after immunization by chloroquine prophylaxis and sporozoites is mediated by preerythrocytic immunity. *Proc Natl Acad Sci U S A*, 110(19):7862–7867, 2013. doi:10.1073/pnas.1220360110.

[14] C. Bisang, C. Weber, J. Inglis, C. Schiffer, W. van Gunsteren, I. Jelesarov, H. Bosshard, and J. Robinson. Stabilization of Type-I .beta.-Turn Conformations in Peptides Containing the NPNA-Repeat Motif of the

Plasmodium falciparum Circumsporozoite Protein by Substituting Proline for (S)-.alpha.-Methylproline. *J Am Chem Soc*, 117(30):7904–7915, 1995. doi:10.1021/ja00135a008.

[15] N. Bowman, S. Congdon, T. Mvalo, J. Patel, V. Escamilla, M. Emch, F. Martinson, I. Hoffman, S. Meshnick, and J. Juliano. Comparative population structure of *Plasmodium falciparum* circumsporozoite protein NANP repeat lengths in Lilongwe, Malawi. *Sci Rep*, 3:1990, 2013. doi:10.1038/srep01990.

[16] C. Busse, I. Czogiel, P. Braun, P. Arndt, and H. Wardemann. Single-cell based high-throughput sequencing of full-length immunoglobulin heavy and light chain genes. *Eur J Immunol*, 44(2):597–603, 2014. doi:10.1002/eji.201343917.

[17] M. Carter, R. Mitchell, P. Meyer Sauteur, D. Kelly, and J. Trück. The Antibody-Secreting Cell Response to Infection: Kinetics and Clinical Applications. *Front Immunol*, 8:630, 2017. doi:10.3389/fimmu.2017.00630.

[18] S. Chaudhury, S. Lyskov, and J. Gray. PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinformatics*, 26(5):689–691, 2010. doi:10.1093/bioinformatics/btq007.

[19] D. Clyde, H. Most, V. McCarthy, and J. Vanderberg. Immunization of man against sporozoite-induced falciparum malaria. *Am J Med Sci*, 266(3):169–77, 1973.

[20] P. Cock, T. Antao, J. Chang, B. Chapman, C. Cox, A. Dalke, I. Friedberg, T. Hamelryck, F. Kauff, B. Wilczynski, and M. de Hoon. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, 25(11):1422–1423, 2009. doi:10.1093/bioinformatics/btp163.

[21] A. Collis, A. Brouwer, and A. Martin. Analysis of the antigen combining site: correlations between length and sequence composition of the hypervariable loops and the nature of the antigen. *J Mol Biol*, 325(2):337–354, 2003. doi:10.1016/S0022-2836(02)01222-6.

[22] M. Corcoran, G. Phad, N. Bernat, C. Stahl-Hennig, M. Sumida, Nand Persson, M. Martin, and G. Karlsson HHedestam. Production of individualized V gene databases reveals high levels of immunoglobulin genetic diversity. *Nature Communications*, 7:13642, 2016. doi:10.1038/ncomms13642.

[23] N. Dauby, C. Kummert, S. Lecomte, C. Liesnard, M. Delforge, C. Donner, and A. Marchant. Primary human cytomegalovirus infection induces the expansion of virus-specific activated and atypical memory B cells. *J Infect Dis.*, 210(8):1275–85, 2014. doi:10.1093/infdis/jiu255.

[24] B. DeKosky, O. Lungu, D. Park, E. Johnson, W. Charab, C. Chrysostomou, D. Kuroda, A. Ellington, G. Ippolito, J. Gray, and G. Georgiou. Large-scale sequence and structural comparisons of human naive and antigen-experienced antibody repertoires. *Proc Natl Acad Sci U S A*, 113(19):E1636–2645, 2016. doi:10.1073/pnas.1525510113.

[25] A. Doedée, N. Kannegieter, K. Öztürk, H. van Loveren, R. Janssen, and A. Buisman. Higher numbers of memory B-cells and Th2-cytokine skewing in high responders to hepatitis B vaccination. *Vaccine*, 34(19), 2016. doi:10.1016/j.vaccine.2015.12.027.

[26] J. Dups, M. Pepper, and I. Cockburn. Antibody and B cell responses to *Plasmodium* sporozoites. *Front Microbiol*, 5:625, 2014. doi:10.3389/fmicb.2014.00625.

[27] R. Edgar. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*, 32(5):1792–1797, 2004. doi:10.1093/nar/gkh340.

[28] Y. Elhanati, Z. Sethna, Q. Marcou, C. Callan, T. Mora, and A. Walczak. Inferring processes underlying B-cell repertoire diversity. *Philos T R Soc B*, 370(1676), 2014. doi:10.1098/rstb.2014.0243.

[29] B. Ellis, P. Haaland, F. Hahne, N. Le Meur, N. Gopalakrishnan, J. Spidlen, and M. Jiang. flowCore: Basic structures for flow cytometry data. *R package version 1.44.0.*, 2017.

[30] C. Fisher, H. Sutton, J. Kaczmarski, H. McNamara, B. Clifton, J. Mitchell, Y. Cai, N. Dups, D'Arcy, M. Singh, A. Chuah, T. Peat, C. Jackson, and I. Cockburn. T-dependent B cell responses to *Plasmodium* induce antibodies that form a high-avidity multivalent complex with the circumsporozoite protein. *PLoS Pathog*, 13 (7):e1006469, 2017. doi:10.1371/journal.ppat.1006469.

[31] F. Fowkes, J. Richards, J. Simpson, and J. Beeson. The relationship between anti-merozoite antibodies and incidence of *Plasmodium falciparum* malaria: A systematic review and meta-analysis. *PLoS Med*, 7(1): e1000218, 2010. doi:10.1371/journal.pmed.1000218.

[32] S. Frevert, U Engelmann, S. Zougbédé, J. Stange, B. Ng, K. Matuschewski, L. Liebes, and H. Yee. Intravital Observation of *Plasmodium berghei* Sporozoite Infection of the Liver. *PLoS Biology*, 3(6):e192, 2005. doi:10.1371/journal.pbio.0030192.

[33] U. Frevert, P. Sinnis, C. Cerami, W. Shreffler, B. Takacs, and V. Nussenzweig. Malaria circumsporozoite protein binds to heparan sulfate proteoglycans associated with the surface membrane of hepatocytes. *J Exp Med*, 177(5):1287–98, 1993. doi:10.1084/jem.177.5.1287.

[34] A. Frosch, O. Odumade, J. Taylor, K. Ireland, G. Ayodo, B. Ondigo, D. Narum, J. Vulule, and C. John. Decrease in Numbers of Naive and Resting B Cells in HIV-Infected Kenyan Adults Leads to a Proportional Increase in Total and *Plasmodium falciparum*-Specific Atypical Memory B Cells. *J Immunol*, 198(12):4629–4638, 2017. doi:10.4049/jimmunol.1600773.

[35] J. Galson, J. Trück, A. Fowler, E. Clutterbuck, M. Münz, V. Cerundolo, C. Reinhard, R. van der Most, A. Pollard, G. Lunter, and D. Kelly. Analysis of B Cell Repertoire Dynamics Following Hepatitis B Vaccination in Humans, and Enrichment of Vaccine-specific Antibody Sequences. *EBioMedicine*, 2(12):2070–9, 2015. doi:10.1016/j.ebiom.2015.11.034.

[36] A. Ghasparian, K. Moehle, A. Linden, and J. Robinson. Crystal structure of an NPNA-repeat motif from the circumsporozoite protein of the malaria parasite *Plasmodium falciparum*. *Chem Commun*, 0(2):174–176, 2006. doi:10.1039/b510812h.

[37] N. Gupta, J. Vander Heiden, M. Uduman, D. Gadala-Maria, G. Yaari, and S. Kleinstein. Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinformatics*, 31(20):3356–8, 2015. doi:10.1093/bioinformatics/btv359.

[38] A. Hagberg, D. Schult, and P. Swart. *Proceedings of the 7th Python in Science Conference (SciPy2008)*, chapter Exploring network structure, dynamics, and function using NetworkX, pages 11–15. Gäel Varoquaux, Travis Vaught, and Jarrod Millman, 2008.

[39] T. Hattori, D. Lai, I. Dementieva, S. Montaño, K. Kurosawa, Y. Zheng, L. Akin, K. Świst Rosowska, A. Grzybowski, A. Koide, K. Krajewski, B. Strahl, N. Kelleher, A. Ruthenburg, and S. Koide. Antigen clasping by two antigen-binding sites of an exceptionally specific antibody for histone methylation. *Proc Natl Acad Sci U S A*, 113(8), 2016. doi:10.1073/pnas.1522691113.

[40] J. Healer, C. Chiu, and D. Hansen. Mechanisms of naturally acquired immunity to *P. falciparum* and approaches to identify merozoite antigen targets. *Parasitology*, 16:1–9, 2017. doi:10.1017/S0031182017001949.

[41] S. Hoffman, C. Oster, C. Plowe, G. Woollett, J. Beier, J. Chulay, R. Wirtz, M. Hollingdale, and M. Mugambi. Naturally acquired antibodies to sporozoites do not prevent malaria: vaccine development implications. *Science*, 237(4815):639–42, 1987. doi:10.1126/science.3299709.

[42] S. Hoffman, L. Goh, T. Luke, I. Schneider, T. Le, D. Doolan, J. Sacci, P. de la Vega, M. Dowler, C. Paul, D. Gordon, J. Stoute, L. Church, M. Sedegah, D. Heppner, W. Ballou, and T. Richie. Protection of humans against malaria by immunization with radiation-attenuated *Plasmodium falciparum* sporozoites. *J Infect Dis*, 185(8):1155–64, 2002. doi:10.1086/339409.

[43] M. Hollingdale, F. Zavala, R. Nussenzweig, and V. Nussenzweig. Antibodies to the protective antigen of *Plasmodium berghei* sporozoites prevent entry into cultured cells. *J Immunol*, 128(4):1929–30, 1982. URL `http://www.jimmunol.org/content/128/4/1929`.

[44] D. Hou, T. Ying, L. Wang, C. Chen, S. Lu, Q. Wang, E. Seeley, J. Xu, X. Xi, T. Li, J. Liu, X. Tang, Z. Zhang, J. Zhou, C. Bai, C. Wang, M. Byrne-Steele, J. Qu, J. Han, and Y. Song. Immune Repertoire Diversity Correlated with Mortality in Avian Influenza A (H7N9) Virus Infected Patients. *Sci Rep*, 6:33843, 2016. doi:10.1038/srep33843.

[45] J. Hunter. Matplotlib: A 2D graphics environment. *Computing In Science & Engineering*, 9(3):90–95, 2007. doi:10.1109/MCSE.2007.55.

[46] Z. Idris, C. Chan, J. Kongere, T. Hall, J. Logedi, J. Gitaka, C. Drakeley, and A. Kaneko. Naturally acquired antibody response to *Plasmodium falciparum* describes heterogeneity in transmission on islands in Lake Victoria. *Sci Rep*, 7(1):9123, 2017. doi:10.1038/s41598-017-09585-4.

[47] K. Imkeller, P. Arndt, H. Wardemann, and C. Busse. sciReptor: analysis of single-cell level immunoglobulin repertoires. *BMC Bioinformatics*, 17:67, 2016. doi:10.1186/s12859-016-0920-1.

[48] K. Jackson, Y. Liu, K. Roskin, J. Glanville, R. Hoh, K. Seo, E. Marshall, T. Gurley, M. Moody, B. Haynes, E. Walter, H. Liao, R. Albrecht, A. García-Sastre, J. Chaparro-Riggers, A. Rajpal, J. Pons, B. Simen, B. Hanczaruk, C. Dekker, J. Laserson, D. Koller, M. Davis, A. Fire, and S. Boyd. Human responses to influenza vaccination show seroconversion signatures and convergent antibody rearrangements. *Cell Host Microbe*, 16(1):105–14, 2014. doi:10.1016/j.chom.2014.05.013.

[49] J. Jeliazkov, A. Sljoka, D. Kuroda, N. Tsuchimura, N. Katoh, K. Tsumoto, and J. Gray. Repertoire analysis of antibody CDR-H3 loops suggests affinity maturation does not typically result in rigidification. *bioRxiv*, 2017. doi:10.1101/230417.

[50] G. Keitany, K. Kim, A. Krishnamurty, B. Hondowicz, W. Hahn, N. Dambrauskas, D. Sather, A. Vaughan, S. Kappe, and M. Pepper. Blood Stage Malaria Disrupts Humoral Immunity to the Pre-erythrocytic Stage Circumsporozoite Protein. *Cell Rep*, 17(12):3193–3205, 2016. doi:10.1016/j.celrep.2016.11.060.

[51] S. Khurana, K. Chung, E. Coyle, A. Meijer, and H. Goldinga. Antigenic Fingerprinting of Antibody Response in Humans following Exposure to Highly Pathogenic H7N7 Avian Influenza Virus: Evidence for Anti-PA-X Antibodies. *J Virol*, 90(20):9383–9393, 2016. doi:10.1128/JVI.01408-16.

[52] I. Kifer, R. Nussinov, and H. Wolfson. GOSSIP: a method for fast and accurate global alignment of protein structures. *Bioinformatics*, 27(7):925–932, 2011. doi:10.1093/bioinformatics/btr044.

[53] E. Krissinel and K. Henrick. Inference of macromolecular assemblies from crystalline state. *J Mol Biol*, 372 (3):774–97, 2007. doi:10.1016/j.jmb.2007.05.022.

[54] M. Krzywinski, J. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S. Jones, , and M. Marra. Circos: An information aesthetic for comparative genomics. *Genome Res*, 19(9):1639–45, 2009. doi:10.1101/gr.092759.109.

[55] T. Kurosaki, K. Kometani, and W. Ise. Memory B cells. *Nat Rev Immunol*, 15(3):149–159, 2015. doi:doi:10.1038/nri3802.

[56] M. Lefranc and G. Lefranc. *The Immunoglobulin FactsBook*. San Diego, CA, Academic Press, 2001.

[57] B. Lell, B. Mordmüller, J. Dejon Agobe, J. Honkpehedji, J. Zinsou, J. Boex Mengue, M. Massinga Loembe, A. Adegnika, J. Held, A. Lalremruata, T. Nguyen, M. Esen, N. KC, A. Ruben, S. Chakravarty, B. Sim, P. Billingsley, E. James, T. Richie, S. Hoffman, and P. Kremsner. Impact of sickle cell trait and naturally acquired immunity on uncomplicated malaria after controlled human malaria infection in adults in Gabon. *Am J Trop Med Hyg.*, 2017. doi:10.4269/ajtmh.17-0343.

[58] D. Lingwood, P. McTamney, H. Yassine, J. Whittle, X. Guo, J. Boyington, C. Wei, and G. Nabel. Structural and genetic basis for development of broadly neutralizing influenza antibodies. *Nature*, 489(7417):566–570, 2012. doi:10.1038/nature11371.

[59] C. MacRaild, J. Richards, R. Anders, and R. Norton. Antibody Recognition of Disordered Antigens. *Structure*, 24(1):148–157, 2016. doi:10.1016/j.str.2015.10.028.

[60] A. Masella, A. Bartram, J. Truszkowski, D. Brown, and J. Neufeld. PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics*, 13:31, 2012. doi:10.1186/1471-2105-13-31.

[61] C. Minici, M. Gounari, R. Übelhart, L. Scarfò, M. Dühren-von Minden, D. Schneider, A. Tasdogan, A. Alkhatib, A. Agathangelidis, S. Ntoufa, N. Chiorazzi, H. Jumaa, K. Stamatopoulos, P. Ghia, and M. Degano. Distinct homotypic B-cell receptor interactions shape the outcome of chronic lymphocytic leukaemia. *Nat Commun.*, 8, 2017. doi:10.1038/ncomms15746.

[62] S. Mishra, R. Nussenzweig, and V. Nussenzweig. Antibodies to *Plasmodium* circumsporozoite protein (CSP) inhibit sporozoite's cell traversal activity. *J Immunol Methods*, 377(1-2):47–52, 2012. doi:10.1016/j.jim.2012.01.009.

[63] S. Moir, J. Ho, A. Malaspina, W. Wang, A. DiPoto, M. O'Shea, G. Roby, S. Kottilil, J. Arthos, M. Proschan, T. Chun, and A. Fauci. Evidence for HIV-associated B cell exhaustion in a dysfunctional memory B cell compartment in HIV-infected viremic individuals. *J Exp Med*, 205(8):1797–1805, 2008. doi:10.1084/jem.20072683.

[64] B. Mordmüller, G. Surat, H. Lagler, S. Chakravarty, A. Ishizuka, A. Lalremruata, M. Gmeiner, J. Campo, M. Esen, A. Ruben, J. Held, C. Calle, J. Mengue, T. Gebru, J. Ibáñez, M. Sulyok, E. James, P. Billingsley, K. Natasha, A. Manoj, T. Murshedkar, A. Gunasekera, A. Eappen, T. Li, R. Stafford, M. Li, P. Felgner, R. Seder, T. Richie, B. Sim, S. Hoffman, and P. Kremsner. Sterile protection against human malaria by chemoattenuated PfSPZ vaccine. *Nature*, 542(7642):445–449, 2017. doi:10.1038/nature21060.

[65] M. Mota, G. Pradel, J. Vanderberg, J. Hafalla, U. Frevert, R. Nussenzweig, V. Nussenzweig, and A. Rodríguez. Migration of *Plasmodium* sporozoites through cells before infection. *Science*, 291(5501):141–4, 2001. doi:10.1126/science.291.5501.141.

[66] M. Muellenbeck, B. Ueberheide, B. Amulic, A. Epp, D. Fenyo, C. Busse, M. Esen, M. Theisen, B. Mordmüller, and H. Wardemann. Atypical and classical memory B cells produce *Plasmodium falciparum* neutralizing antibodies. *J Exp Med*, 210(2):389–399, 2013. doi:10.1084/jem.20121970.

[67] K. Murphy. *Janeway's Immunobiology*. Number 8th Edition. Garland Science, 2012.

[68] R. Murugan, K. Imkeller, C. Busse, and H. Wardemann. Direct high-throughput amplification and sequencing of immunoglobulin genes from single human B cells. *Eur J Immunol*, 45(9):2698–700, 2015. doi:10.1002/eji.201545526.

[69] R. Murugan, L. Buchauer, G. Triller, C. Kreschel, G. Costa, G. Pidelaserra Marti, K. Imkeller, C. Busse, S. Chakravarty, B. Sim, S. Hoffman, E. Levashina, P. Kremsner, B. Mordmüller, T. Höfer, and W. H. Clonal selection drives protective memory B cell responses in controlled human malaria infection. *Science Immunology, in press*, 2018.

[70] W. Nahrendorf, A. Scholzen, E. Bijker, A. Teirlinck, G. Bastiaens, R. Schats, C. Hermsen, L. Visser, J. Langhorne, and R. Sauerwein. Memory B-Cell and Antibody Responses Induced by *Plasmodium falciparum* Sporozoite Immunization. *J Infect Dis*, 210(12):1981–1990, 2014. doi:10.1093/infdis/jiu354.

[71] D. R. J. North B, Lehmann A. A new clustering of antibody CDR loop conformations. *J Mol Biol*, 406(2): 228–256, 2011. doi:10.1016/j.jmb.2010.10.030.

[72] R. Nussenzweig, J. Vanderberg, H. Most, and C. Orton. Protective immunity produced by the injection of x-irradiated sporozoites of *Plasmodium berghei*. *Nature*, 216(5111):160–2, 1967.

[73] V. Offeddu, V. Thathy, K. Marsh, and K. Matuschewski. Naturally acquired immune responses against *Plasmodium falciparum* sporozoites and liver infection. *Int J Parasitol.*, 42(6):535–48, 2012. doi:10.1016/j.ijpara.2012.03.011.

[74] F. Osier, G. Fegan, S. Polley, L. Murungi, F. Verra, K. Tetteh, B. Lowe, T. Mwangi, P. Bull, A. Thomas, D. Cavanagh, J. McBride, D. Lanar, M. Mackinnon, D. Conway, and K. Marsh. Breadth and Magnitude of Antibody Responses to Multiple *Plasmodium falciparum* Merozoite Antigens Are Associated with Protection from Clinical Malaria. *Infect Immun*, 76(5):2240–8, 2008. doi:10.1128/IAI.01585-07.

[75] D. Oyen, J. Torres, U. Wille-Reece, C. Ockenhouse, D. Emerling, J. Glanville, W. Volkmuth, Y. Flores-Garcia, F. Zavala, A. Ward, C. King, and I. Wilson. Structural basis for antibody recognition of the NANP repeats in *Plasmodium falciparum* circumsporozoite protein. *Proc Natl Acad Sci U S A*, 114(48):E10438–E10445, 2017. doi:10.1073/pnas.1715812114.

[76] P. Parameswaran, Y. Liu, K. Roskin, K. Jackson, V. Dixit, J. Lee, K. Artiles, S. Zompi, M. Vargas, B. Simen, B. Hanczaruk, K. McGowan, M. Tariq, N. Pourmand, D. Koller, A. Balmaseda, S. Boyd, E. Harris, and A. Fire. Convergent antibody signatures in human dengue. *Cell Host Microbe*, 13(6):691–700, 2013. doi:10.1016/j.chom.2013.05.008.

[77] PATH Malaria Vaccine Initiative. http://www.malariavaccine.org/malaria-and-vaccines/vaccine-development/life-cycle-malaria-parasite. 2017.

[78] G. Pidelaserra Marti. Internship report: Diversity of the human antibody repertoire in malaria immunity. Technical report, University of Heidelberg, Faculty of Biosciences, 2016.

[79] M. Plassmeyer, K. Reiter, S. R. Jr, S. Kotova, P. D. Smith, D. E. Hurt, B. House, X. Zou, Y. Zhang, M. Hickman, O. Uchime, R. Herrera, V. Nguyen, J. Glen, J. Lebowitz, A. J. Jin, L. H. Miller, N. J. MacDonald, Y. Wu, and D. L. Narum. Structure of the *Plasmodium falciparum* circumsporozoite protein, a leading malaria vaccine candidate. *JBC*, 284(39), 2009. doi:10.1074/jbc.M109.013706.

[80] S. Portugal, C. Tipton, H. Sohn, Y. Kone, J. Wang, S. Li, J. Skinner, K. Virtaneva, D. Sturdevant, S. Porcella, O. Doumbo, S. Doumbo, K. Kayentao, A. Ongoiba, B. Traore, I. Sanz, S. Pierce, and P. Crompton. Malaria-associated atypical memory B cells exhibit markedly reduced B cell receptor signaling and effector function. *Elife*, 4, 2015. doi:10.7554/eLife.07218.

[81] P. Potocnjak, N. Yoshida, R. Nussenzweig, and V. Nussenzweig. Monovalent fragments (Fab) of monoclonal antibodies to a sporozoite surface antigen (Pb44) protect mice against malarial infection. *J Exp Med*, 151(6): 1504–13, 1980. doi:10.1084/jem.151.6.1504.

[82] P. Potocnjak, F. Zavala, R. Nussenzweig, and V. Nussenzweig. Inhibition of idiotype–anti-idiotype interaction for detection of a parasite antigen: a new immunoassay. *Science*, 215(4540):1637–1639, 1982. doi:10.1126/science.6122269.

[83] G. Pradel and U. Frevert. Malaria sporozoites actively enter and pass through rat Kupffer cells prior to hepatocyte invasion. *Hepatology*, 33(5):1154–1165, 2001. doi:10.1053/jhep.2001.24237.

[84] T. Ramaraj, T. Angel, E. Dratz, A. Jesaitis, and B. Mumey. Antigen-antibody interface properties: composition, residue interactions, and features of 53 non-redundant structures. *Biochim Biophys Acta.*, 1824(3):520–32, 2012. doi:10.1016/j.bbapap.2011.12.007.

[85] J. Rini, R. Stanfield, E. Stura, P. Salinas, A. Profy, and I. Wilson. Crystal structure of a human immunodeficiency virus type 1 neutralizing antibody, 50.1, in complex with its V3 loop peptide antigen. *Proc Natl Acad Sci U S A.*, 90(13):6325–9, 1993. URL `http://www.pnas.org/content/90/13/6325.long`.

[86] G. Robin, Y. Sato, D. Desplancq, N. Rochel, E. Weiss, and P. Martineau. Restricted diversity of antigen binding residues of antibodies revealed by computational alanine scanning of 227 antibody-antigen complexes. *J Mol Biol.*, 426(22):3729–43, 2014. doi:10.1016/j.jmb.2014.08.013.

[87] R. Rosenberg, R. Wirtz, I. Schneider, and R. Burge. An estimation of the number of malaria sporozoites ejected by a feeding mosquito. *Trans R Soc Trop Med Hyg.*, 84(2):209–12, 1990.

[88] B. Roy, R. Neumann, O. Snir, R. Iversen, G. Sandve, K. Lundin, and L. Sollid. High-throughput single-cell analysis of b cell receptor usage among autoantigen-specific plasma cells in celiac disease. *J Immunol.*, 199 (2):782–791, 2017. doi:10.4049/jimmunol.1700169.

[89] RTS,S Clinical Trials Partnership. A Phase 3 Trial of RTS,S/AS01 Malaria Vaccine in African Infants. *N Engl J Med*, 367, 2012. doi:10.1056/NEJMoa1208394.

[90] 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(9988):31–45, 2015. doi:10.1016/S0140-6736(15)60721-8.

[91] E. Saphire, P. Parren, R. Pantophlet, M. Zwick, G. Morris, P. Rudd, R. Dwek, R. Stanfield, D. Burton, and I. Wilson. Crystal structure of a neutralizing human IGG against HIV-1: a template for vaccine design. *Science*, 293(5532):1155–9, 2001. doi:10.1126/science.1061692.

[92] J. Scheid, H. Mouquet, B. Ueberheide, R. Diskin, F. Klein, T. Oliveira, J. Pietzsch, D. Fenyo, A. Abadir, K. Velinzon, A. Hurley, S. Myung, F. Boulad, P. Poignard, D. Burton, F. Pereyra, D. Ho, B. Walker, M. Seaman, P. Bjorkman, B. Chait, and M. Nussenzweig. Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science*, 333(6049):1633–1637, 2011. doi:10.1126/science.1207227.

[93] A. Schmidt, M. Therkelsen, S. Stewart, T. Kepler, H. Liao, M. Moody, B. Haynes, and S. Harrison. Viral Receptor-Binding Site Antibodies with Diverse Germline Origins. *Cell*, 161(5):1026–1034, 2015. doi:10.1016/j.cell.2015.04.028.

[94] L. Schofield, J. Villaquiran, A. Ferreira, H. Schellekens, R. Nussenzweig, and V. Nussenzweig. Gamma interferon, CD8+ T cells and antibodies required for immunity to malaria sporozoites. *Nature*, 330(6149):664–6, 1987. doi:10.1038/330664a0.

[95] R. Seder, L. Chang, M. Enama, K. Zephir, U. Sarwar, I. Gordon, L. Holman, E. James, P. Billingsley, A. Gunasekera, A. Richman, S. Chakravarty, A. Manoj, S. Velmurugan, M. Li, A. Ruben, T. Li, A. Eappen, R. Stafford, S. Plummer, C. Hendel, L. Novik, P. Costner, F. Mendoza, J. Saunders, M. Nason, J. Richardson, J. Murphy, S. Davidson, T. Richie, M. Sedegah, A. Sutamihardja, G. Fahle, K. Lyke, M. Laurens, M. Roederer, K. Tewari, J. Epstein, B. Sim, J. Ledgerwood, B. Graham, and S. V. . S. T. Hoffman. Protection against malaria by intravenous immunization with a nonreplicating sporozoite vaccine. *Science*, 341(6152):1359–65, 2013. doi:10.1126/science.1241800.

[96] R. Shinnakasu, T. Inoue, K. Kometani, S. Moriyama, Y. Adachi, M. Nakayama, Y. Takahashi, H. Fukuyama, T. Okada, and T. Kurosaki. Regulated selection of germinal-center cells into the memory B cell compartment. *Nat Immunol*, 17(7):861–9, 2016. doi:10.1038/ni.3460.

[97] P. Sinnis, P. Clavijo, D. Fenyö, B. Chait, C. Cerami, and V. Nussenzweig. Structural and functional properties of region II-plus of the malaria circumsporozoite protein. *J Exp Med*, 180(1):297–306, 1994. doi:10.1084/jem.180.1.297.

[98] K. Smith, J. Muther, A. Duke, E. McKee, N. Zheng, P. Wilson, and J. James. Fully human monoclonal antibodies from antibody secreting cells after vaccination with Pneumovax®23 are serotype specific and facilitate opsonophagocytosis. *Immunobiology*, 218(5):745–54, 2013. doi:10.1016/j.imbio.2012.08.278.

[99] H. Song, X. Nie, S. Basu, M. Singh, and J. Cerny. Regulation of VH gene repertoire and somatic mutation in germinal centre B cells by passively administered antibody. *Immunology*, 98(2):258–66, 1999. doi:10.1046/j.1365-2567.1999.00874.x.

[100] R. Sullivan, C. Kim, M. Fontana, M. Feeney, P. Jagannathan, M. Boyle, C. Drakeley, I. Ssewanyana, F. Nankya, H. Mayanja-Kizza, G. Dorsey, and B. Greenhouse. FCRL5 Delineates Functionally Impaired Memory B Cells Associated with *Plasmodium falciparum* Exposure. *PLoS Pathog*, 11(5):e1004894, 2015. doi:10.1371/journal.ppat.1004894.

[101] K. Tewari, B. Flynn, S. Boscardin, K. Kastenmueller, A. Salazar, C. Anderson, V. Soundarapandian, A. Ahumada, T. Keler, S. Hoffman, M. Nussenzweig, R. Steinman, and R. Seder. Poly(I:C) is an effective adjuvant for antibody and multi-functional CD4+ T cell responses to *Plasmodium falciparum* circumsporozoite protein (CSP) and anti-DEC-CSP in non human primates. *Vaccine*, 28:7256–7266, 2010. doi:10.1016/j.vaccine.2010.08.098.

[102] C. Thomson, K. Little, D. Reason, and J. Schrader. Somatic diversity in CDR3 loops allows single V-genes to encode innate immunological memories for multiple pathogens. *J Immunol*, 186(4):2291–2298, 2011. doi:10.4049/jimmunol.0904092.

[103] T. Tiller, E. Meffre, S. Yurasov, M. Tsuiji, M. Nussenzweig, and H. Wardemann. Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. *J Immunol Methods.*, 329(1-2), 2008. doi:10.1016/j.jim.2007.09.017.

[104] G. Triller, S. Scally, G. Costa, M. Pissarev, C. Kreschel, A. Bosch, E. Marois, B. Sack, R. Murugan, A. Salman, C. Janse, S. Khan, S. Kappe, A. Adegnika, B. Mordmüller, E. Levashina, J. Julien, and H. Wardemann. Natural parasite exposure induces protective human anti-malarial antibodies. *Immunity*, 47(6):1197–1209.e10, 2017. doi:10.1016/j.immuni.2017.11.007.

[105] J. Trück, M. Ramasamy, J. Galson, R. Rance, J. Parkhill, G. Lunter, A. Pollard, and D. Kelly. Identification of antigen-specific B cell receptor sequences using public repertoire analysis. *J Immunol*, 194(1):252–261, 2015. doi:10.4049/jimmunol.1401405.

[106] S. van der Walt, S. Colbert, and G. Varoquaux. The NumPy Array: A Structure for Efficient Numerical Computation. *Computing in Science & Engineering*, 13:22–30, 2011. doi:10.1109/MCSE.2011.37.

[107] J. Vander Heiden, G. Yaari, M. Uduman, J. Stern, K. O'Connor, D. Hafler, F. Vigneault, and S. Kleinstein. pRESTO: a toolkit for processing high-throughput sequencing raw reads of lymphocyte receptor repertoires. *Bioinformatics*, 30(13):1930–1932, 2014. doi:10.1093/bioinformatics/btu138.

[108] J. Vanderberg and U. Frevert. Intravital microscopy demonstrating antibody-mediated immobilisation of *Plasmodium berghei* sporozoites injected into skin by mosquitoes. *Int J Parasitol*, 34(9):991–6, 2004. doi:10.1016/j.ijpara.2004.05.005.

[109] G. Victora and M. Nussenzweig. Germinal centers. *Annu Rev Immunol*, 30:429–457, 2012. doi:10.1146/annurev-immunol-020711-075032.

[110] C. Vollmers, R. Sit, J. Weinstein, C. Dekker, and S. Quake. Genetic measurement of memory B-cell recall using antibody repertoire sequencing. *Proc Natl Acad Sci U S A*, 110(33):13463–13468, 2013. doi:10.1073/pnas.1312146110.

[111] C. Wang, Y. Liu, M. Cavanagh, S. Le Saux, Q. Qi, K. Roskin, T. Looney, J. Lee, V. Dixit, C. Dekker, G. Swan, J. Goronzy, and S. Boyd. B-cell repertoire responses to varicella-zoster vaccination in human identical twins. *Proc Natl Acad Sci U S A*, 112(2):500–5, 2015. doi:10.1073/pnas.1415875112.

[112] H. Wardemann, S. Yurasov, A. Schaefer, J. Young, E. Meffre, and M. Nussenzweig. Predominant autoantibody production by early human B cell precursors. *Science*, 301(5638):1374–7, 2003. doi:10.1126/science.1086907.

[113] M. Wataya, T. Sano, N. Kamikawaji, T. Tana, K. Yamamoto, and T. Sasazuki. Comparative analysis of HLA restriction and cytokine production in hepatitis B surface antigen-specific T cells from low- and high-antibody responders in vaccinated humans. *J Hum Genet*, 46(4):197–206, 2001. doi:10.1007/s100380170089.

[114] C. Watson, K. Steinberg, J. Huddleston, R. Warren, M. Malig, J. Schein, A. Willsey, J. Joy, J. Scott, T. Graves, R. Wilson, R. Holt, E. Eichler, and F. Breden. Complete haplotype sequence of the human immunoglobulin

heavy-chain variable, diversity, and joining genes and characterization of allelic and copy-number variation. *Am J Hum Genet*, 92(4):530–546, 2013. doi:10.1016/j.ajhg.2013.03.004.

[115] B. F. Watson CT1. The immunoglobulin heavy chain locus: genetic variation, missing data, and implications for human disease. *Genes Immun*, 13(5):363–373, 2012. doi:10.1038/gene.2012.12.

[116] G. Weber, M. White, A. Babakhanyan, P. Sumba, J. Vulule, D. Ely, C. John, E. Angov, D. Lanar, S. Dutta, D. Narum, A. Horii, T ant Cowman, J. Beeson, J. Smith, J. Kazura, and A. Dent. Sero-catalytic and Antibody Acquisition Models to Estimate Differing Malaria Transmission Intensities in Western Kenya. *Sci Rep*, 7(1), 2017. doi:10.1038/s41598-017-17084-9.

[117] H. Webster, A. Brown, C. Chuenchitra, B. Permpanich, and J. Pipithkul. Characterization of antibodies to sporozoites in *Plasmodium falciparum* malaria and correlation with protection. *J Clin Microbiol*, 26(5):923–927, 1988.

[118] D. Weese, M. Holtgrewe, and K. Reinert. RazerS 3: faster, fully sensitive read mapping. *Bioinformatics*, 28 (20):2592–2599, 2012. doi:10.1093/bioinformatics/bts505.

[119] F. Weisel, G. Zuccarino-Catania, M. Chikina, and M. Shlomchik. A Temporal Switch in the Germinal Center Determines Differential Output of Memory B and Plasma Cells. *Immunity*, 44(1):116–130, 2016. doi:10.1016/j.immuni.2015.12.004.

[120] B. Weitzner, J. Jeliazkov, S. Lyskov, N. Marze, D. Kuroda, R. Frick, J. Adolf-Bryfogle, N. Biswas, R. J. Dunbrack, and J. Gray. Modeling and docking of antibody structures with Rosetta. *Nat Protoc*, 12(2): 401–416, 2017. doi:10.1038/nprot.2016.180.

[121] M. White, B. P, A. Olotu, J. T. Griffin, E. M. Riley, K. E. Kester, C. F. Ockenhouse, and A. C. Ghani. The relationship between RTS,S vaccine-induced antibodies, CD4(+) T cell responses and protection against *Plasmodium falciparum* infection. *PLoS One*, 8(4):e61395, 2013. doi:10.1371/journal.pone.0061395.

[122] D. Wilson, R. Wirtz, and B. Finlay. Recognition of phage-expressed peptides containing Asx-Pro sequences by monoclonal antibodies produced against *Plasmodium falciparum* circumsporozoite protein. *Protein Engineering, Design and Selection*, 10(5):531–40, 1997. doi:10.1093/protein/10.5.531.

[123] T. Wu and E. Kabat. An analysis of the sequences of the variable regions of bence jones proteins and myeloma light chains and their implications for antibody complementarity. *J Exp Med*, 132(2):211–50, 1970. doi:10.1084/jem.132.2.211.

[124] R. Xu, J. Krause, R. McBride, J. Paulson, J. J. Crowe, and I. Wilson. A recurring motif for antibody recognition of the receptor-binding site of influenza hemagglutinin. *Nat Struct Mol Biol*, 20(3):363–370, 2013. doi:10.1038/nsmb.2500.

[125] G. Yaari, J. Vander Heiden, M. Uduman, D. Gadala-Maria, N. Gupta, J. Stern, K. O'Connor, D. Hafler, U. Lasserson, F. Vigneault, and K. S. Models of somatic hypermutation targeting and substitution based on synonymous mutations from high-throughput immunoglobulin sequencing data. *Frontiers in Immunology*, 4: 358, 2013. doi:10.3389/fimmu.2013.00358.

[126] N. Ye, Jand Ma, T. Madden, and J. Ostell. IgBLAST: an immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Res*, 41(Web Server issue):W34–W40, Jul 2013. doi:10.1093/nar/gkt382.

[127] T. Ying, P. Prabakaran, L. Du, W. Shi, Y. Feng, Y. Wang, L. Wang, W. Li, S. Jiang, D. Dimitrov, and T. Zhou. Junctional and allele-specific residues are critical for MERS-CoV neutralization by an exceptionally potent germline-like antibody. *Nat Commun*, 6:8223, 2015. doi:10.1038/ncomms9223.

[128] N. Yoshida, P. Potocnjak, V. Nussenzweig, and R. Nussenzweig. Biosynthesis of Pb44, the protective antigen of sporozoites of *Plasmodium berghei*. *J Exp Med*, 154(4):1225, 1981. doi:10.1084/jem.154.4.1225.

[129] M. Yuda and T. Ishino. Liver invasion by malarial parasites–how do malarial parasites break through the host barrier? *Cell Microbiol*, 6(12):1119–25, 2004. doi:10.1111/j.1462-5822.2004.00474.x.

[130] F. Zavala, A. Cochrane, E. Nardin, R. Nussenzweig, and V. Nussenzweig. Circumsporozoite proteins of malaria parasites contain a single immunodominant region with two or more identical epitopes. *J Exp Med*, 157(6): 1947–57, 1983. doi:10.1084/jem.157.6.1947.

[131] D. R. Zerbino, P. Achuthan, W. Akanni, M. Amode, D. Barrell, J. Bhai, K. Billis, C. Cummins, A. Gall, C. G. Girón, L. Gil, L. Gordon, L. Haggerty, E. Haskell, T. Hourlier, O. G. Izuogu, S. H. Janacek, T. Juettemann, J. K. To, M. R. Laird, I. Lavidas, Z. Liu, J. E. Loveland, T. Maurel, W. McLaren, B. Moore, J. Mudge, D. N.

Murphy, V. Newman, M. Nuhn, D. Ogeh, C. K. Ong, A. Parker, M. Patricio, H. S. Riat, H. Schuilenburg, D. Sheppard, H. Sparrow, K. Taylor, A. Thormann, A. Vullo, B. Walts, A. Zadissa, A. Frankish, S. E. Hunt, M. Kostadima, N. Langridge, F. J. Martin, M. Muffato, E. Perry, M. Ruffier, D. M. Staines, S. J. Trevanion, B. L. Aken, F. Cunningham, A. Yates, and P. Flicek. Ensembl 2018. *Nucleic Acids Res*, 46(D1):D754–D761, 2018. doi:10.1093/nar/gkx1098.

[132] Y. Zhang, M. Meyer-Hermann, L. George, M. Figge, M. Khan, M. Goodall, S. Young, A. Reynolds, F. Falciani, A. Waisman, C. Notley, M. Ehrenstein, M. Kosco-Vilbois, and K. Toellner. Germinal center B cells govern their own fate via antibody feedback. *J Exp Med*, 210(3):457–64, 2013. doi:10.1084/jem.20120150.

12 Acknowledgements

The work leading to the completion of this thesis would not have been possible without significant help of many people:

I would like to thank Prof. Dr. Hedda Wardemann for giving me the possibility to perform this work in the Department of B cell Immunology at the German Cancer Research Center in Heidelberg. I would also like to thank Prof. Dr. Thomas Höfer, Prof. Dr. Rebecca Wade and Dr. Aleksandra Walzcak for their advice and scientific input during the thesis advisory committee meetings. During my PhD work I was a student in the International Max Planck Research School for Infection Biology and in the Helmholtz International Graduate School for Cancer Research. The research internship at the Scripps Research Institute was funded by EMBO.

My thesis work involved a lot of programming and setup of bioinformatics tools. I would like to thank Dr. Peter Arndt and Dr. Irina Czogiel for helping me getting started with the analysis software at the Max Planck Institute for Molecular Genetics. I would like to thank Prof. William Schief for giving me the possibility to visit his lab at the Scripps Research Institute and improve my antibody modelling skills. I thank Dr. Jared Adolf-Bryfogle and Dr. Sebastian Rämisch for supervising me during this internship and sharing their expertise. Setting up the antibody structure modelling software at the DKFZ would not have been possible without the help of Dr. Nikolaus Kepper and Erik Bernstein.

The PBMC samples from Gabonese donors were collected during the LaCHMI trial at the Centre de Recherches Médicales de Lambaréné. It was a pleasure for me to be part of the LaCHMI team and I would like to thank especially Aurore Bouyoukou Hounkpatin, Dr. Bertrand Lell, Dr. Marguerite Massinga Loembe and Prof. Dr. Benjamin Mordmüller.

I performed the experiments with *Pf* sporozoites in the Vector Biology department at the Max Planck Institute in Berlin. I would like to thank the whole team for helping me with the preparation and for welcoming me in their lab, especially Cornelia Kreschel, Dr. Yara Reis, Dr. Giulia Costa and Prof. Dr. Elena Levashina.

Finally, I would like to thank all group members of of the Division of B cell Immunology at the DKFZ in Heidelberg for creating a very nice and stimulating atmosphere in the laboratory. I thank Gemma Pidelaserra Martí, Dorien Foster and Claudia Winter for their help with experiments and Dr. Gianna Triller and Rajagopal Murugan for beeing my malaria research community at DKFZ. My sincere thanks go to Dr. Christian Busse, who was always there to help me throughout the complete course of my thesis and to answer all of my questions.

13 Appendix

13.1 FACS analysis during LaCHMI trial

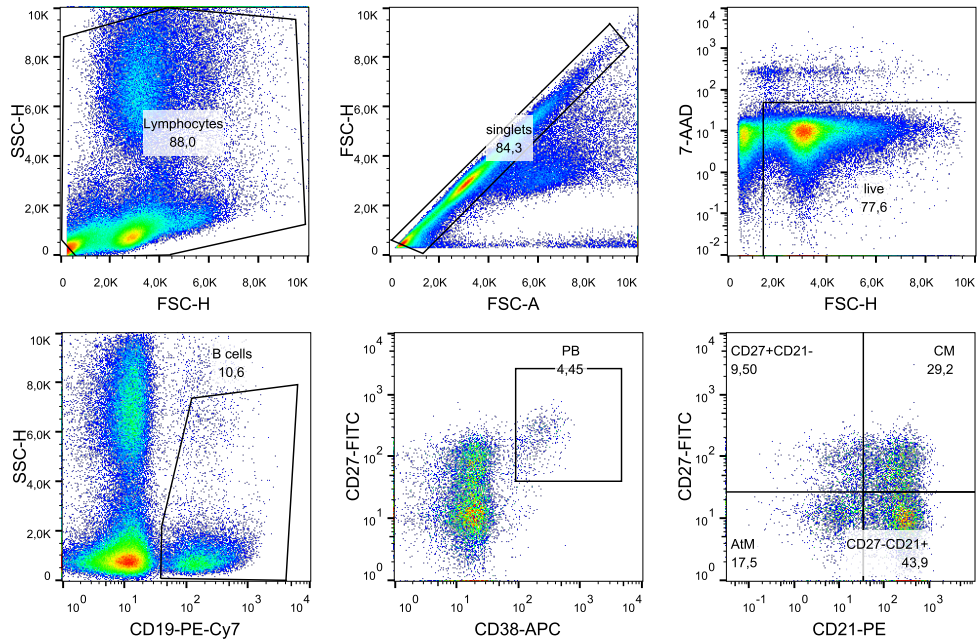
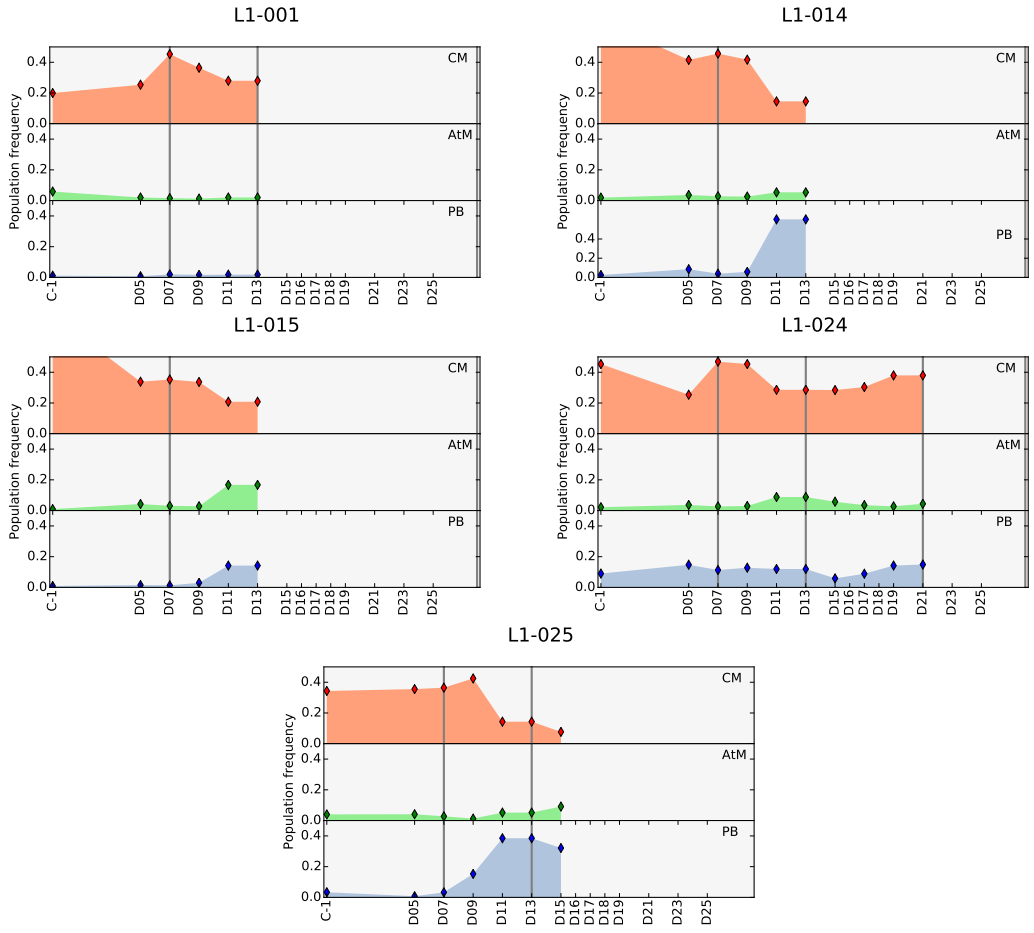


Figure 24. Gating strategy for measurement of frequencies of three B cell sub-populations: plasmablast (PB), atypical memory B cells (AtM) and classical memory B cells (CM).

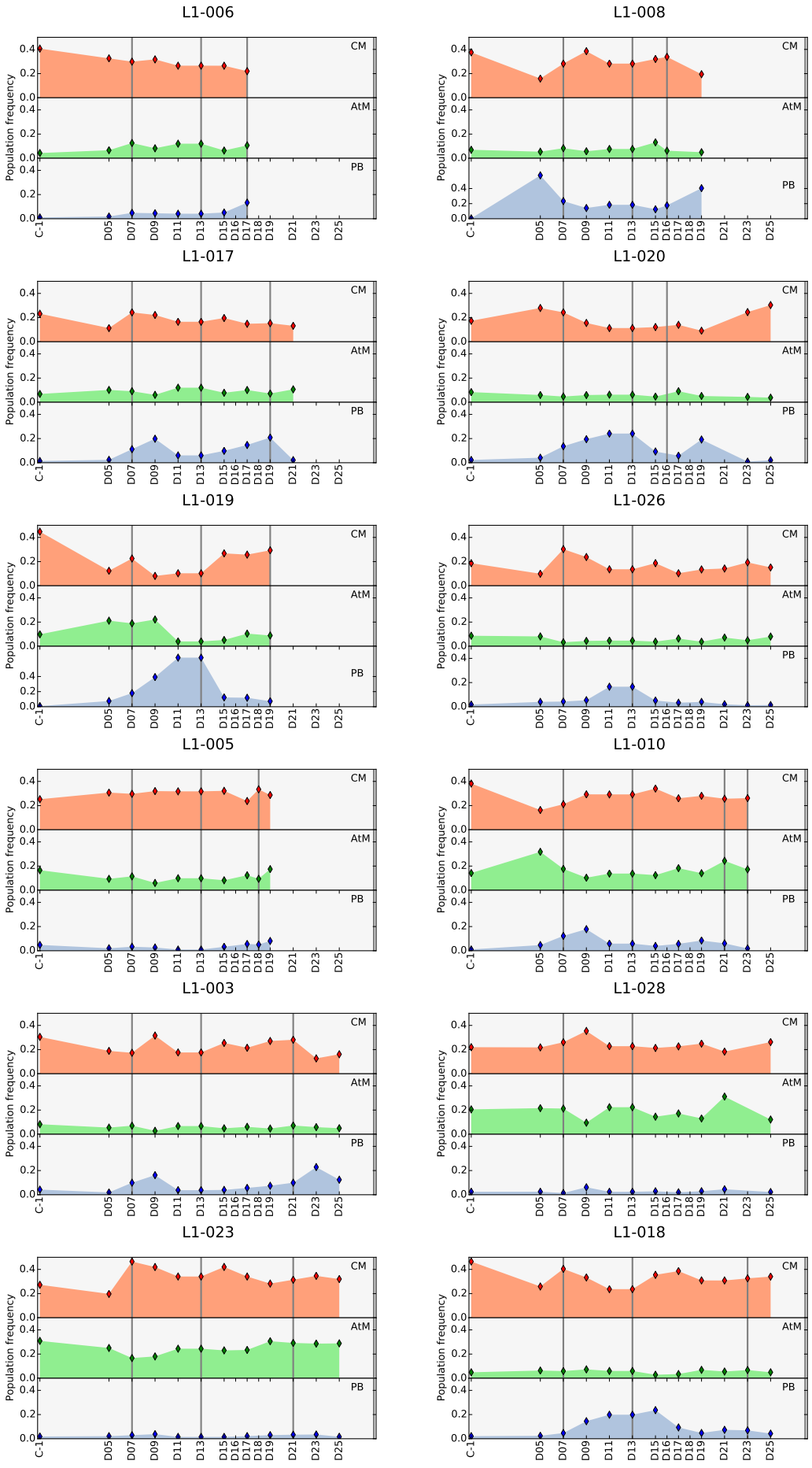
13.1.1 B cell population frequencies in the different LaCHMI participants

The following panels depict the frequencies of classical memory B cells (CM), atypical memory B cells (AtM) and plasmablasts (PB) within the complete CD19+ B cell population. The donors are classified according to their background (European, Gabonese) and according to whether developed blood stage parasitemia or not. Grey vertical lines indicate timepoints at which 20 mL of blood were sampled and PBMCs cryo-preserved.

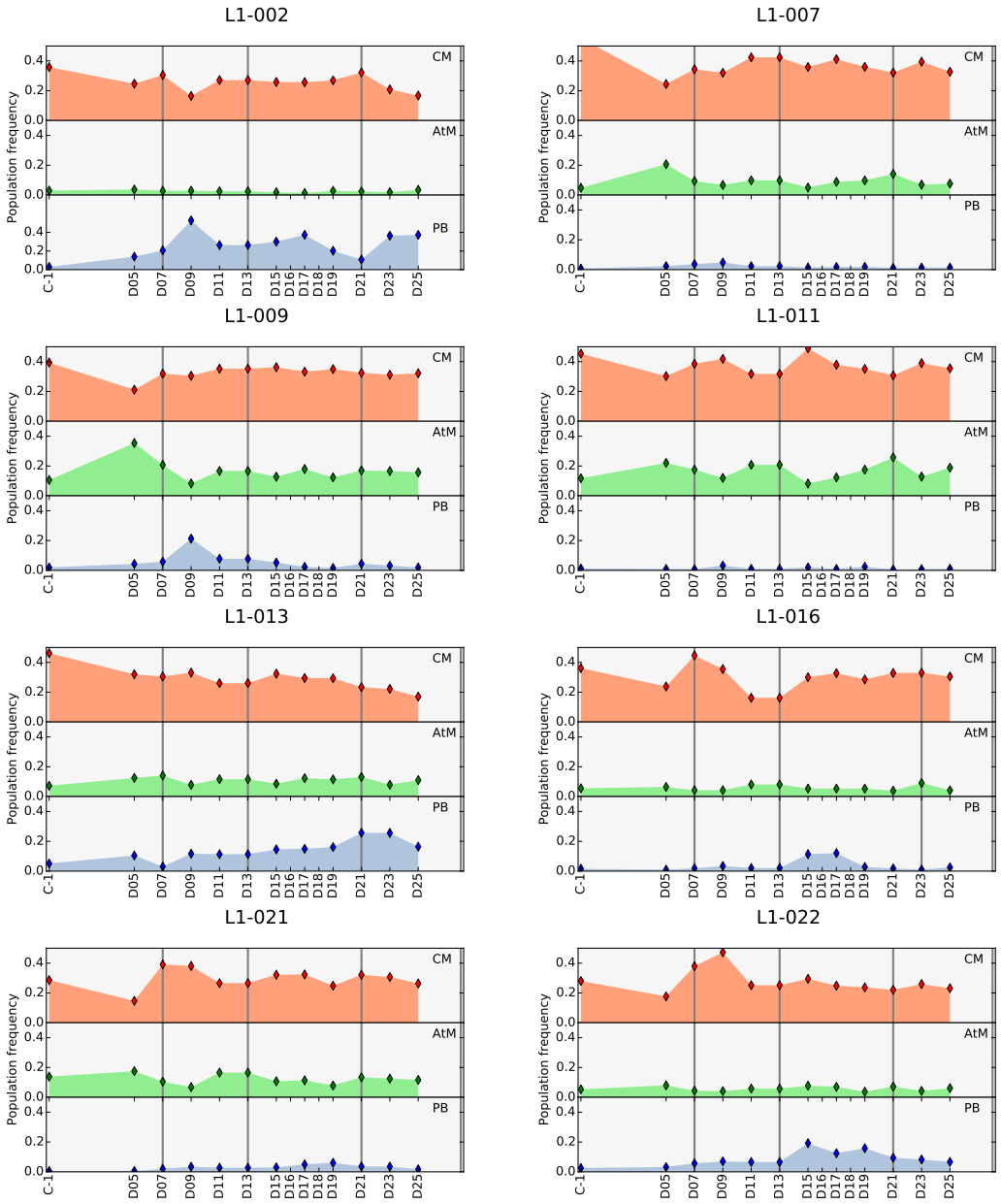
Naive European LaCHMI donors



Gabonese LaCHMI donors with parasitemia



Gabonese LaCHMI donors without parasitemia



1210 variants IgH	IGHV3-33*01 (H.31, H.50, H.56) — D — IGHJ3*02
1210_GL	QVQLVESGGGVVQPGRSLRLSCAASGFTFS**S**YGMH**W**VRQAPGKGLEWVA**V**IWYDGS**N**KYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARVRDSSDYYGDAFDIWGQGTMVTVSS
1210	QVQLVESGGGVVQPGRSLRLSCAASGFTFS**N**YGMH**W**VRQAPGKGLEWVA**V**IWYDGS**K**KYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARVRDSSDYYGDAFDIWGQGTMVTVSS
1210_H.V50_I^{mut}	QVQLVESGGGVVQPGRSLRLSCAASGFTFS**N**YGMH**W**VRQAPGKGLEWVA**I**IWYDGS**K**KYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARVRDSSDYYGDAFDIWGQGTMVTVSS
1210_H.N31_S^{rev}	QVQLVESGGGVVQPGRSLRLSCAASGFTFS**S**YGMH**W**VRQAPGKGLEWVA**V**IWYDGS**K**KYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARVRDSSDYYGDAFDIWGQGTMVTVSS
1210_H.K57_N^{rev}	QVQLVESGGGVVQPGRSLRLSCAASGFTFS**N**YGMH**W**VRQAPGKGLEWVA**V**IWYDGS**N**KYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARVRDSSDYYGDAFDIWGQGTMVTVSS
1210_K.N93_S^{rev}	QVQLVESGGGVVQPGRSLRLSCAASGFTFS**N**YGMH**W**VRQAPGKGLEWVA**V**IWYDGS**K**KYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARVRDSSDYYGDAFDIWGQGTMVTVSS
1210_H.K57_N^{rev}_K.N93_S^{rev} Named 1210_NS	QVQLVESGGGVVQPGRSLRLSCAASGFTFS**N**YGMH**W**VRQAPGKGLEWVA**V**IWYDGS**N**KYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARVRDSSDYYGDAFDIWGQGTMVTVSS
1210_H.D104Y_N^{mut}_K.N92_Y^{mut} Named 1210_YY	QVQLVESGGGVVQPGRSLRLSCAASGFTFS**N**YGMH**W**VRQAPGKGLEWVA**V**IWYDGS**K**KYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARVRDSS**Y**YYGDAFDIWGQGTMVTVSS

1210 variants Igκ	IGKV1-5*03 (K.93) — IGKJ1*01
1210_GL	DIQMTQSPSTLSASVGDRVTITC**R**ASQSISSWLAWYQQKPGKAPKLLIYKAS**S**LESGVP**S**RFSGSGSGTEFTLTISSLQPDDFATYYCQQYN**S**YWTFGQGTKVEIK
1210	DIQMTQSPSTLSASVGDRVTITC**R**ASQSISSWLAWYQQKPGKAPKLLIYKAS**N**LESGVP**L**RFSGSGSGTEFTLTISSLQPDDFATYYCQQYN**K**YWTFGQGTKVEIK
1210_H.V50_I^{mut}	DIQMTQSPSTLSASVGDRVTITC**R**ASQSISSWLAWYQQKPGKAPKLLIYKAS**N**LESGVP**L**RFSGSGSGTEFTLTISSLQPDDFATYYCQQYN**K**YWTFGQGTKVEIK
1210_H.N31_S^{rev}	DIQMTQSPSTLSASVGDRVTITC**R**ASQSISSWLAWYQQKPGKAPKLLIYKAS**N**LESGVP**L**RFSGSGSGTEFTLTISSLQPDDFATYYCQQYN**K**YWTFGQGTKVEIK
1210_H.K57_N^{rev}	DIQMTQSPSTLSASVGDRVTITC**R**ASQSISSWLAWYQQKPGKAPKLLIYKAS**N**LESGVP**L**RFSGSGSGTEFTLTISSLQPDDFATYYCQQYN**K**YWTFGQGTKVEIK
1210_K.N93_S^{rev}	DIQMTQSPSTLSASVGDRVTITC**R**ASQSISSWLAWYQQKPGKAPKLLIYKAS**N**LESGVP**L**RFSGSGSGTEFTLTISSLQPDDFATYYCQQYN**S**YWTFGQGTKVEIK
1210_H.K57_N^{rev}_K.N93_S^{rev} Named 1210_NS	DIQMTQSPSTLSASVGDRVTITC**R**ASQSISSWLAWYQQKPGKAPKLLIYKAS**N**LESGVP**L**RFSGSGSGTEFTLTISSLQPDDFATYYCQQYN**S**YWTFGQGTKVEIK
1210_H.D104Y_N^{mut}_K.N92_Y^{mut} Named 1210_YY	DIQMTQSPSTLSASVGDRVTITC**R**ASQSISSWLAWYQQKPGKAPKLLIYKAS**N**LESGVP**L**RFSGSGSGTEFTLTISSLQPDDFATYYCQQY**Y**YWTFGQGTKVEIK

2163 variants IgH	IGHV3-33*01 (H.31, H.50, H.56) — D — IGHJ4*02
2163	QVQLVESGGGVVQPGRSLRLSCAASGFTFS**S**YGMH**W**VRQAPGKGLEWVA**V**IWYDGS**N**KYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARVQTTTGGGSCCPFDYWGQGTLVTVSS
2163_H.V50_I^{mut}	QVQLVESGGGVVQPGRSLRLSCAASGFTFS**S**YGMH**W**VRQAPGKGLEWVA**I**IWYDGS**N**KYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARVQTTTGGGSCCPFDYWGQGTLVTVSS
2163_H.S31_N^{mut}	QVQLVESGGGVVQPGRSLRLSCAASGFTFS**N**YGMH**W**VRQAPGKGLEWVA**V**IWYDGS**N**KYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARVQTTTGGGSCCPFDYWGQGTLVTVSS
2163_H.N57_K^{mut}	QVQLVESGGGVVQPGRSLRLSCAASGFTFS**S**YGMH**W**VRQAPGKGLEWVA**V**IWYDGS**K**KYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARVQTTTGGGSCCPFDYWGQGTLVTVSS
2163_H.N57_K^{mut}_K.S93_N^{mut} Named 2163_KN	QVQLVESGGGVVQPGRSLRLSCAASGFTFS**S**YGMH**W**VRQAPGKGLEWVA**V**IWYDGS**K**KYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARVQTTTGGGSCCPFDYWGQGTLVTVSS

2163 variants Igκ	IGKV1-5*03 (K.93) — IGKJ1*01
2163	DIQMTQSPSSLSASVGDRVTITC**R**ASQSISSWLAWYQQKPGKAPKLLIYKASSLESGVPSRFSGSGSGTEFTLTISSLQPDDFATYYCQQYN**S**YWTFGQGTKVEIK
2163_H.V50_I^{mut}	DIQMTQSPSSLSASVGDRVTITC**R**ASQSISSWLAWYQQKPGKAPKLLIYKASSLESGVPSRFSGSGSGTEFTLTISSLQPDDFATYYCQQYN**S**YWTFGQGTKVEIK
2163_H.S31_N^{mut}	DIQMTQSPSSLSASVGDRVTITC**R**ASQSISSWLAWYQQKPGKAPKLLIYKASSLESGVPSRFSGSGSGTEFTLTISSLQPDDFATYYCQQYN**S**YWTFGQGTKVEIK
2163_H.N57_K^{mut}	DIQMTQSPSSLSASVGDRVTITC**R**ASQSISSWLAWYQQKPGKAPKLLIYKASSLESGVPSRFSGSGSGTEFTLTISSLQPDDFATYYCQQYN**S**YWTFGQGTKVEIK
2163_H.N57_K^{mut}_K.S93_N^{mut} Named 2163_KN	DIQMTQSPSSLSASVGDRVTITC**R**ASQSISSWLAWYQQKPGKAPKLLIYKASSLESGVPSRFSGSGSGTEFTLTISSLQPDDFATYYCQQYN**K**YWTFGQGTKVEIK

grey: CDRs | yellow: mutations at positions H.31, H.50, H.56 and K.93 | blue: mutations at other positions | pink: mutations introducing sterical hindrance

13.3 Primer sequences

13.3.1 PCR primers for generation of primary amplicons

First PCR forward primers

hIGHV-1/7-066-fw	ACAGGTGCCCACTCCCAGGTGCAG
hIGHV-1/7-017-fw	ATGGACTGGACCTGGAG
hIGHV-1/7-041-fw	TCCTCTTTGTGGTGGCAGCAGC
hIGHV-2-035-fw	TCCACGCTCCTGCTRCTGAC
hIGHV-3-066-fw	AAGGTGTCCAGTGTGARGTGCAG
hIGHV-3-057-fw	TAAAAGGTGTCCAGTGT
hIGHV-4/6-066-fw	CCCAGATGGGTCCTGTCCCAGGTGCAG
hIGHV-4-022-fw	ATGAAACACCTGTGGTTCTTCC
hIGHV-5-066-fw	CAAGGAGTCTGTTCCGAGGTGCAG
hIGKV-1-060-fw	ATGAGGSTCCCYGCTCAGCTGCTGG
hIGKV-3-049-fw	CTCTTCCTCCTGCTACTCTGGCTCCCAG
hIGKV-4-049-fw	ATTTCTCTGTTGCTCTGGATCTCTG
hIGLV-1-068-fw	GGTCCTGGGCCCAGTCTGTGCTG
hIGLV-2-068-fw	GGTCCTGGGCCCAGTCTGCCCTG
hIGLV-3-068-fw	GCTCTGTGACCTCCTATGAGCTG
hIGLV-4/5-068-fw	GGTCTCTCTCSCAGCYTGTGCTG
hIGLV-6-068-fw	GTTCTTGGGCCAATTTTATGCTG
hIGLV-7-068-fw	GGTCCAATTCYCAGGCTGTGGTG
hIGLV-8-083-fw	GAGTGGATTCTCAGACTGTGGTG

First PCR reverse primers

hIGHA-111-rv	GTCCGCTTTCGCTCCAGGTCACACT
hIGHE-140-rv	AAGGTCATAGTTGTCCCGTTGAGG
hIGHG-137-rv	GGAAGGTGTGCACGCCGCTGGTC
hIGHM-082-rv	GGAAGGAAGTCCTGTGCGAGGC
hIGKC-172-rv	GTGCTGTCCTTGCTGTCCTGCT
hIGLC-057-rv	CACCAGTGTGGCCTTGTTGGCTTG

Second PCR forward primers

hIGHV-pan-080-fw	AGGTGCAGCTGCTGGAGTCKGG
hIGKV-pan-fw	ATGACCCAGWCTCCABYCWCCCTG
hIGLV-pa1-fw	CAGYCTGYSCTGACTCA
hIGLV-pa2-fw	TCCTATGAGCTGACWCAG

Second PCR reverse primers

hIGHA-076-rv	GGAAGAAGCCCTGGACCAGGC
hIGHE-070-rv	CCAGGCAGCCCAGAGTCACGG
hIGHG-074-rv	AGTCCTTGACCAGGCAGCCC
hIGHM-031-rv	GGGAATTCTCACAGGAGACGA
hIGKC-032-rv	AACTGCTCATCAGATGGCGGG
hIGLC-026-rv	TCAGAGGAGGGYGGGAACAGAGTG

13.3.2 Primers for specific PCR and subsequent cloning

Specific primers for heavy chain

5' AgeI VH1	CTGCAACCGGTGTACATTCCCAGGTGCAGCTGGTGCAG
5' AgeI VH1-18	CTGCAACCGGTGTACATTCCCAGGTTCAGCTGGTGCAG
5' AgeI VH1-24	CTGCAACCGGTGTACATTCCCAGGTCCAGCTGGTACAG
5' AgeI VH1/5	CTGCAACCGGTGTACATTCCGAGGTGCAGCTGGTGCAG
5' AgeI VH3	CTGCAACCGGTGTACATTCTGAGGTGCAGCTGGTGGAG
5' AgeI VH3-9	CTGCAACCGGTGTACATTCTGAAGTGCAGCTGGTGGAG
5' AgeI VH3-23	CTGCAACCGGTGTACATTCTGAGGTGCAGCTGTTGGAG
5' AgeI VH3-33	CTGCAACCGGTGTACATTCTCAGGTGCAGCTGGTGGAG
5' AgeI VH4	CTGCAACCGGTGTACATTCCCAGGTGCAGCTGCAGGAG
5' AgeI VH4-34	CTGCAACCGGTGTACATTCCCAGGTGCAGCTACAGCAGTG
5' AgeI VH4-39	CTGCAACCGGTGTACATTCCCAGCTGCAGCTGCAGGAG
5' AgeI VH6-1	CTGCAACCGGTGTACATTCCCAGGTACAGCTGCAGCAG
5' AgeI VH7	CTGCAACCGGTGTACATTCTCAGGTGCAGCTGGTGCAATCTGG
3' SalI JH1/2/4/5	TGCGAAGTCGACGCTGAGGAGACGGTGACCAG
3' SalI JH3	TGCGAAGTCGACGCTGAAGAGACGGTGACCATTG
3' SalI JH6	TGCGAAGTCGACGCTGAGGAGACGGTGACCGTG

Specific primers for kappa chain

5' AgeI Vk1-5	CTGCAACCGGTGTACATTCTGACATCCAGATGACCCAGTC
5' AgeI Vk1-5	TTGTGCTGCAACCGGTGTACATTCAGACATCCAGTTGACCCAGTCT
5' AgeI Vk1-6	CTGCAACCGGTGTACATTCTGCCATCCAGATGACCCAGTC
5' AgeI Vk1-13	CTGCAACCGGTGTACATTCTGCCATCCAGTTGACCCAGTC
5' AgeI Vk1D-43	CTGCAACCGGTGTACATTGTGCCATCCGGATGACCCAGTC
5' AgeI Vk2-24	CTGCAACCGGTGTACATGGGGATATTGTGATGACCCAGAC
5' AgeI Vk2-28	CTGCAACCGGTGTACATGGGGATATTGTGATGACTCAGTC
5' AgeI Vk2-30	CTGCAACCGGTGTACATGGGGATGTTGTGATGACTCAGTC
5' AgeI Vk3-11	TTGTGCTGCAACCGGTGTACATTCAGAAATTGTGTTGACACAGTC
5' AgeI Vk3-15	CTGCAACCGGTGTACATTCAGAAATAGTGATGACGCAGTC
5' AgeI Vk3-20	TTGTGCTGCAACCGGTGTACATTCAGAAATTGTGTTGACGCAGTCT
5' AgeI Vk4-1	CTGCAACCGGTGTACATTCGGACATCGTGATGACCCAGTC
3' BsiWI Jk1/4	GCCACCGTACGTTTGATYTCCACCTTGGTC
3' BsiWI Jk2	GCCACCGTACGTTTGATCTCCAGCTTGGTC
3' BsiWI Jk3	GCCACCGTACGTTTGATATCCACTTTGGTC
3' BsiWI Jk5	GCCACCGTACGTTTAATCTCCAGTCGTGTC

Specific primers for lambda chain

5' AgeI Vl1	CTGCTACCGGTTCCTGGGCCCAGTCTGTGCTGACKCAG
5' AgeI Vl2	CTGCTACCGGTTCCTGGGCCCAGTCTGCCCTGACTCAG
5' AgeI Vl3	CTGCTACCGGTTCTGTGACCTCCTATGAGCTGACWCAG
5' AgeI Vl4/5	CTGCTACCGGTTCTCTCTCSCAGCYTGTGCTGACTCA
5' AgeI Vl6	CTGCTACCGGTTCTTGGGCCAATTTTATGCTGACTCAG
5' AgeI Vl7/8	CTGCTACCGGTTCCAATTCYCAGRCTGTGGTGACYCAG
hCl-040-XhoI	CTCCTCACTCGAGGGYGGGAACAGAGTG

13.3.3 Primers for insert check

hIGHG-084-Rv	GTTCGGGGAAGTAGTCCTTGAC
hIGKC-172-Rv	GTGCTGTCCTTGCTGTCCTGCT
hIGLC-057-Rv	CACCAGTGTGGCCTTGTTGGCTTG
Absense	GCTTCGTTAGAACGCGGCTAC

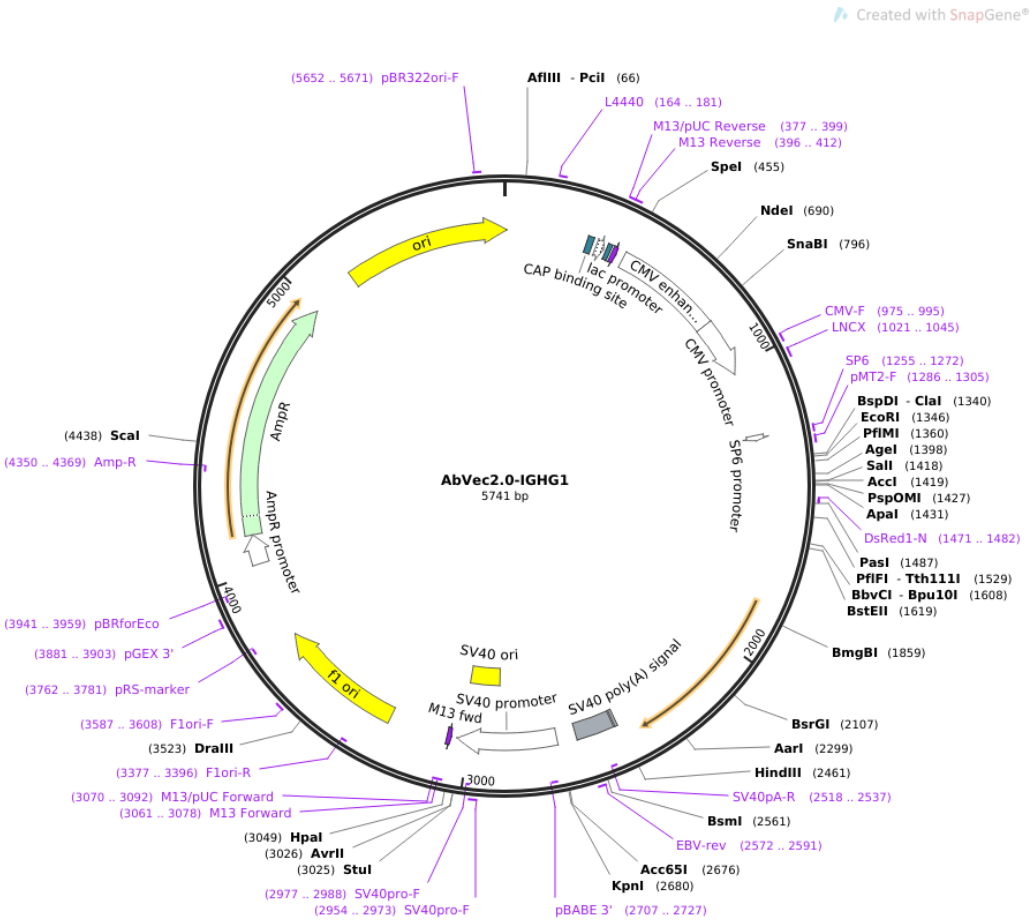
13.3.4 Primers for *IGHV* genotyping

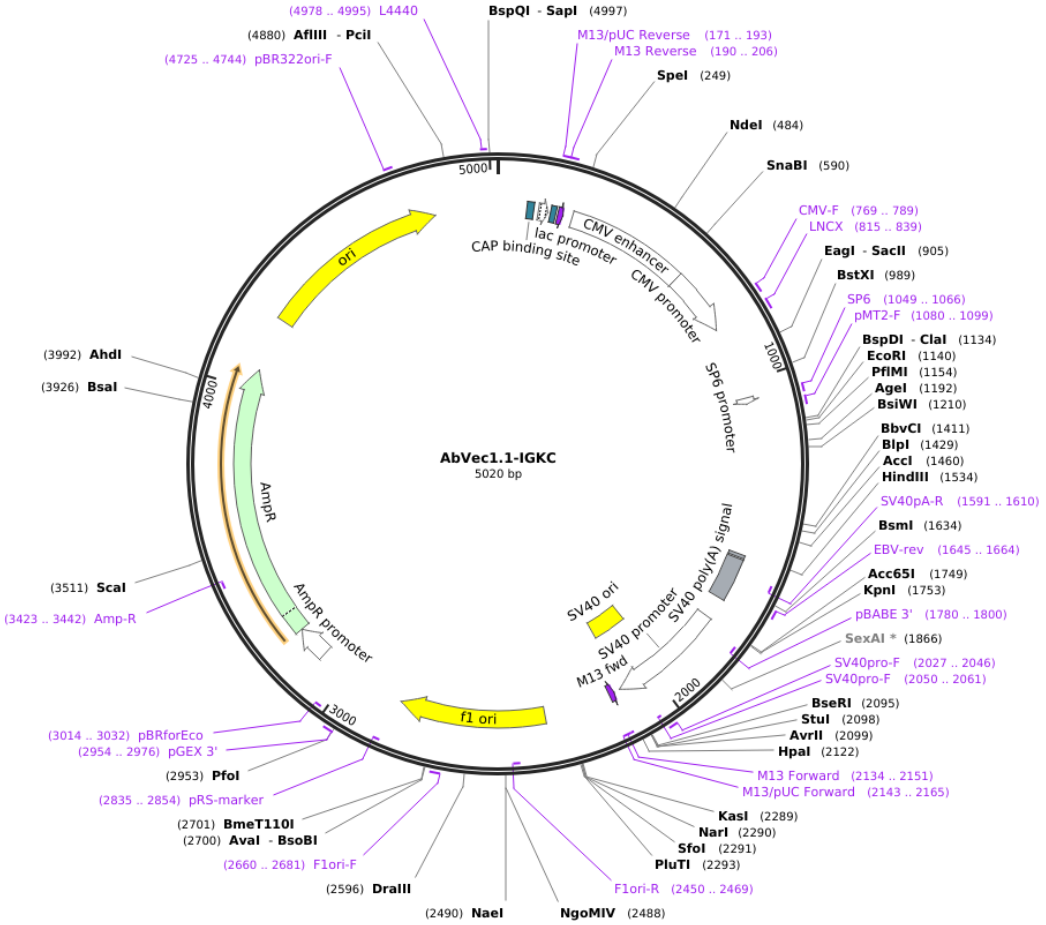
IGHV3-30_079_Fw_01	ACAGGTAGTCCACACACAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_02	ACCAACACCACAGTCACAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_03	ACCTGATCTCGAACCTCAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_04	ACGAGTAGTCTGGAGTCAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_05	ACGAGTTGACCATCCTCAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_06	ACGTACAGCACTACCTCAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_07	ACGTACTCAGACAGTCCAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_08	ACGTAGAGGAGACTCACAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_09	ACGTAGTCCTCTCTGTCAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_10	ACTCGATCCAGTTGTCCAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_11	ACTCGTAGCTGTGAACCAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_12	ACTCGTCTACTGTGTCCAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_13	ACTGAGACGAACACGACAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_14	AGAGCTACCAGTTCACCAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_15	AGAGCTCAAGTGGTTCCAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_16	AGGACAGTAGTGGTCACAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_17	AGGTTCACTGAGAGGACAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_18	AGGTTCCAGACTCACACAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_19	AGTCCTCTGTCAAGGTCAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_20	AGTCCTGTTCTCTGGACAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_21	AGTCCTTCGATCTCCTCAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_22	AGTGTCACCTTCGAGACAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_23	CAACGTCTTGAGAGTGCAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_24	CAAGAGGACAGTTCTGCAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_25	CACTACAGGTGTTCCACAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_26	TGTCAGGACTTGGAGTCAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_27	TGTGGACTAGAGGTAGCAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_079_Fw_28	TGTGGTAGACACGAGACAGGTGCAGCTGGTGGAGTCTGGGG
IGHV3-30_348_Rv	TCTCTCGCACAGTAATACACAGCCGTGTCC
3' SalI JH1/2/4/5	TGCGAAGTCGACGCTGAGGAGACGGTGACCAG
3' SalI JH3	TGCGAAGTCGACGCTGAAGAGACGGTGACCATTG
3' SalI JH6	TGCGAAGTCGACGCTGAGGAGACGGTGACCGTG

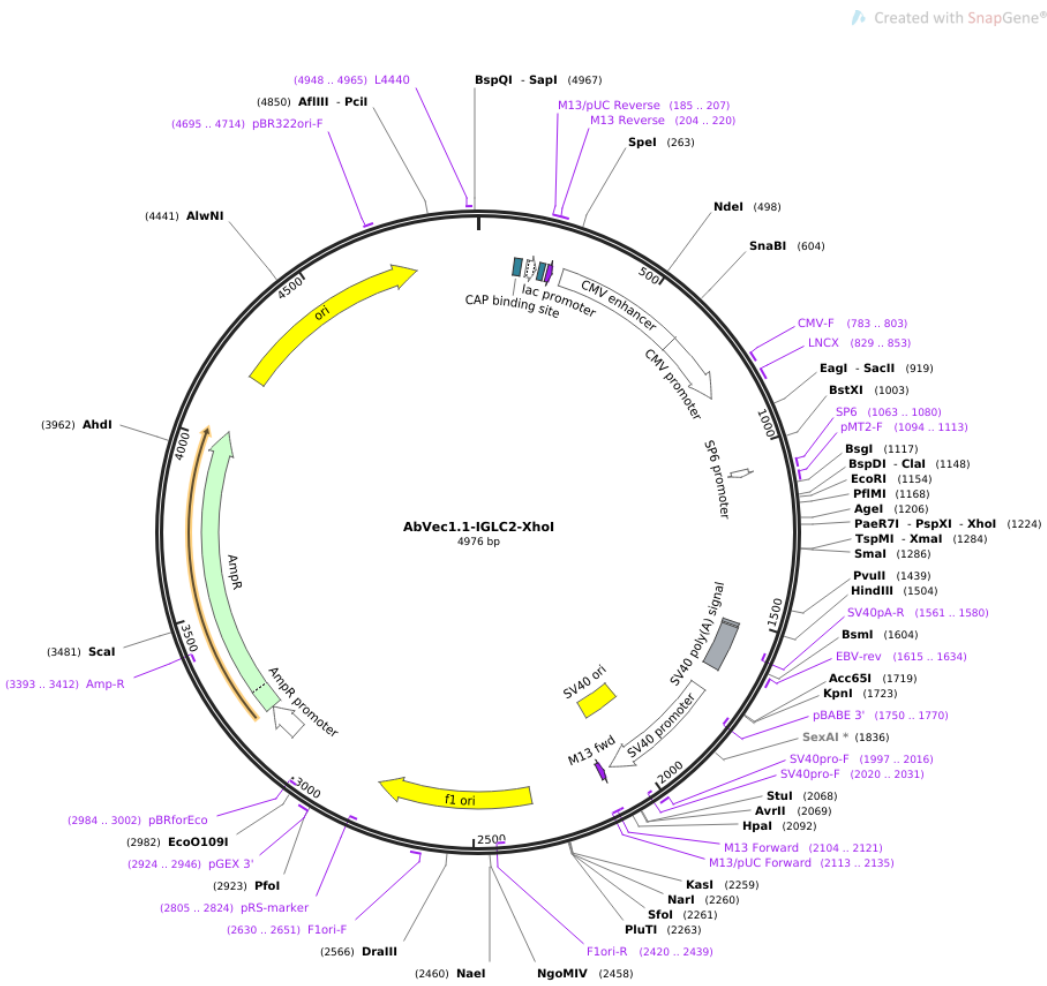
13.3.5 Primers for site directed mutagenesis

Mutant	Forward primer	Reverse primer
D07g1x1349_T84A	CATCCTGAGAgccGAGGACACGG	TTCATTTGCAGATACAGTGAGTTCTTG
D07xxl1349_N31K	CATCGGAAGTaagTATCTATACTGG	TTGGAGCTGCTTCCAGAAC
D07g1x1349_Y58A	GTGGGTTTCAgccATTAGTAGTAGTAGTAGTACCATATACTAC	TCCAGCCCCTTCCCTGGA
D07xxl1349_Y34A	TAATTATCTAgccTGGTACCAGCAGC	CTTCCGATGTTGGAGCTG
D0972h2219_W52R	GGCAGTTATAaggTATGATGGAAGTAAAAAATACTATG	ACCCACTCCAGCCCCTTG
D0972h2219_Y53A	AGTTATATGGgctGATGGAAGTAAAAAATACTATGCAGACTCCGTGAAG	GCCACCCACTCCAGCCCC
D0972k2219_K50E	CCTGATCTATgagGCGTCTCGTTTAG	AGCTTAGGGGCTTTCCCT
D0972k2219_W35A	TATTAGTAGCgcgTTGGCCTGGTATCAGCAG	CTCTGACTGGCCCGGCAA
D0972h2140_W52A	GGCAATTATAgcgTATGATGGAAGTAAGAAATACTATGCAGAC	ACCCACTCCAGCCCCTTG
D0972h2140_W52R	GGCAATTATAaggTATGATGGAAGTAAGAAATACTATG	ACCCACTCCAGCCCCTTG
D0972h2140_W52S	GGCAATTATAagcTATGATGGAAGTAAGAAATACTATGCAG	ACCCACTCCAGCCCCTTG
D0972h2140_Y59A	AAGTAAGAAAgccTATGCAGACTCCG	CCATCATACCATATAATTGCC
D0972k2140_F97Y	GTATAAGAGTtacTGGACGTTCG	TGTTGGCAGTAATAAGTTG
D0972k2140_W35A	TATTAGTAGCgcgTTGGCCTGGTATCAGCAG	CTCTGACTGGCCCGGCAA
D0972h2290_W52R	GGCAGTTATAaggTATGATGGAAGTAATAAATACTATG	ACCCACTCCAGCCCCTTG
D0972h2290_W52S	GGCAGTTATAtcgTATGATGGAAGTAATAAATACTATG	ACCCACTCCAGCCCCTTG
D0972h2290_N57K	TGATGGAAGTaagAAATACTATGC	TACCATATAACTGCCACC
D0972h2290_V50I	GTGGGTGGCAattATATGGTATGATGG	TCCAGCCCCTTGCCTGGA
D0773h1239_R52W	AGGTTTCATTtggAGCAAAGCTTATGGTGG	ACCCACTCCAGCCCCTTC
D0972h2163_W52S	GGCAGTTATAtcgTATGATGGAAGTAATAAATACTATG	ACCCACTCCAGCCCCTTG
D0773h1210_W52A	GGCAGTTATAaggTATGATGGAAGTAAGAAATAC	ACCCACTCCAGCCCCTTG
D0773h1210_K57N	TGATGGAAGTaacAAATACTATGCAG	TACCATATAACTGCCACC
D0773h1210_V50I	GTGGGTGGCAattATATGGTATGATGG	TCCAGCCCCTTGCCTGGA
D0773h1210_D104Y	TGATAGTAGTtatTATTATGGTGATGCTTTTG	CGAACTCTCGCACAGTAATAC
D0773h1210_N92Y	CCAACAATATtatAATTATTGGACG	CAGTAATAAGTTGCAAAATCATC
D0773h1210_Y59A	AAGTAAGAAAgccTATGCAGACTCCG	CCATCATACCATATAACTGCC
D0773h1210_KN93S	ACAATATAATagtTATTGGACGTTCG	TGGCAGTAATAAGTTGCAAAATC
2163_KS93N	ACAGTATAATaatTATTGGACGTTC	TGGCAGTAATAAGTTGCAAAATC
D0773h1210_HN31S	CACCTTCAGTagcTATGGCATGC	AATCCAGACGCTGCACAG
2163_HN31S	CACCTTCAGTaacTATGGCATGC	AATCCAGACGCTGCACAG
D0773h1210_Y53A	AGTTATATGGgctGATGGAAGTAAGAAATACTATGCAGACTCCGTGAAGG	GCCACCCACTCCAGCCCC
D0773h1210_V50A	GTGGGTGGCAgctATATGGTATGATG	TCCAGCCCCTTGCCTGGA
D0773h1210_V50F_W52R	acggTATGATGGAAGTAAGAAATACTATGCAGACTCCG	ataaaTGCCACCCACTCCAGCCC
2163_V50F_W52RF	acggTATGATGGAAGTAATAAATACTATGCAGACTCCGTG	ataaaTGCCACCCACTCCAGCCC
D0773h1450_Y59A	TGGTGACACAgctTACGCGGACTC	CTACTGCCACTAATACCTG
D0773h1450_S35H	CTACGGCATGcatTGGGTCCGCC	TTGCTAAAGCTGAATCCAG
1450_H.SGS52WYDF	cgatAGTGGTGACACATATTAC	taacaAATACCTGAGACCCACTC

13.4 Expression vector maps






Created with SnapGene®
AbVec1.1-IGLC2-XhoI
4976 bp
BspQI - SapI (4967)
(4948 .. 4965) L4440
(4850) AflIII - PciI
(4695 .. 4714) pBR322ori-F
M13/pUC Reverse (185 .. 207)
M13 Reverse (204 .. 220)
SpeI (263)
NdeI (498)
SnaBI (604)
(4441) AlwNI
CMV-F (783 .. 803)
LNCX (829 .. 853)
EagI - SacII (919)
BstXI (1003)
SP6 (1063 .. 1080)
pMT2-F (1094 .. 1113)
BsgI (1117)
BspDI - ClaI (1148)
EcoRI (1154)
PflMI (1168)
AgeI (1206)
PaeR7I - PspXI - XhoI (1224)
TspMI - XmaI (1284)
SmaI (1286)
PvuII (1439)
HindIII (1504)
SV40pA-R (1561 .. 1580)
BsmI (1604)
EBV-rev (1615 .. 1634)
Acc65I (1719)
KpnI (1723)
pBABE 3' (1750 .. 1770)
SexAI * (1836)
SV40pro-F (1997 .. 2016)
SV40pro-F (2020 .. 2031)
StuI (2068)
AvrII (2069)
HpaI (2092)
M13 Forward (2104 .. 2121)
M13/pUC Forward (2113 .. 2135)
KasI (2259)
NarI (2260)
SfoI (2261)
PluTI (2263)
F1ori-R (2420 .. 2439)
NgoMIV (2458)
(2460) NaeI
(2566) DraIII
(2630 .. 2651) F1ori-F
(2805 .. 2824) pRS-marker
(2923) PfoI
(2924 .. 2946) pGEX 3'
(2982) EcoO109I
(2984 .. 3002) pBRforEco
(3393 .. 3412) Amp-R
(3481) ScaI
(3962) AhdI
ori
lac promoter
CAP binding site
CMV enhancer
CMV promoter
SP6 promoter
SV40 poly(A) signal
SV40 promoter
SV40 ori
M13 fwd
f1 ori
AmpR promoter
AmpR
500
1000
1500
2000
2500
3000
3500
4000
4500

13.5 List of materials

ELISA capture antibody	
Goat anti-human IgG, Fcgamma	Jackson, ImmunoResearch Laboratories, West Grove, PA, USA
ELISA secondary antibodies	
Goat anti-human IgG, Fcgamma (HRP conjugated)	Jackson, ImmunoResearch Laboratories, West Grove, PA, USA
Goat anti-human IgM, Fc5mu (HRP conjugated)	Jackson, ImmunoResearch Laboratories, West Grove, PA, USA
Goat anti-human IgA, alpha Chain (HRP conjugated)	Jackson, ImmunoResearch Laboratories, West Grove, PA, USA
ELISA standard and control antibodies	
2A10 (chimeric antibody)	Triller et al. 2017
Human IgG1, kappa	Sigma Aldrich Chemie GmbH, Steinheim, Germany
mGO53 (non-poly-reactive antibody)	Wardemann et al. 2003
FACS analysis antibodies and reagents	
7-AAD	Invitrogen GmbH, Karlsruhe, Germany
Mouse anti-human CD19 (BV786-conjugated, clone SJ25C1))	BD Biosciences GmbH, Heidelberg, Germany
Mouse anti-human CD19 (PE-Cy7-conjugated, clone SJ25C1))	BD Biosciences GmbH, Heidelberg, Germany
Mouse anti-human CD21 (PE-Cy7-conjugated, clone Bu32)	BioLegend GmbH, Fell, Germany
Mouse anti-human CD27 (PE-conjugated, Nr. 555441)	BD Biosciences GmbH, Heidelberg, Germany
Mouse anti-human IgD (V500-conjugated, clone 1A6-2)	BD Biosciences GmbH, Heidelberg, Germany
Mouse anti-human IgG (APC-H7-conjugated, clone 1A6-2)	BD Biosciences GmbH, Heidelberg, Germany
Mouse anti-human CD20 (APC-H7-conjugated, clone G18-145)	BD Biosciences GmbH, Heidelberg, Germany
Mouse anti-human CD20 (BV711-conjugated, clone 2H7)	BD Biosciences GmbH, Heidelberg, Germany
Mouse anti-human CD38 (FITC-conjugated, Nr. 555459)	BD Biosciences GmbH, Heidelberg, Germany
Mouse anti-human CD138 (BV421-conjugated, clone MI15)	BD Biosciences GmbH, Heidelberg, Germany
Antigens	
CSP	Kind gift of Dr. Kim Lee Sim, Protein Potential LLC, Rockville, MD, USA
Δ-NCSP	Kind gift of Dr. Silvia Boscardin, Department of Parasitology, University of São Paulo, Brazil
$NANP_5$, $NANP_{10}$	Alpha Diagnostic Intl. Inc., Texas, USA
$NANP_3$	PSL GmbH, Heidelberg, Germany
Bacteria and Media	Clonetech Inc., Palo Alto, CA, USA
E.coli DH10B	
LB agar	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Lysogeny broth (LB)	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Terrific broth (TB)	Life Technologies GmbH, Karlsruhe, Germany
Chemicals, buffers and solutions	
5x loading buffer gel electrophoresis	
50x TAE buffer	
ABTS self-made buffer	0.1 M citric acid, 0.2 M disodium phosphate
ABTS tablets	Roche Diagnostics GmbH, Mannheim, Germany
Acetic acid (CH_3COOH)	Sigma Aldrich Chemie GmbH, Steinheim, Germany
Ammonium chloride (NH_4Cl_2)	Sigma Aldrich Chemie GmbH, Steinheim, Germany
Ammonium sulfate ($(NH_4)_2SO_4$)	Sigma Aldrich Chemie GmbH, Steinheim, Germany
Bovine serum albumin fraction V (BSA)	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Calcium chloride ($CaCl_2$)	Sigma Aldrich Chemie GmbH, Steinheim, Germany
Citric acid ($C_6H_8O_7$)	Sigma Aldrich Chemie GmbH, Steinheim, Germany
Cytochalasin D	Sigma Aldrich Chemie GmbH, Steinheim, Germany
Dextran-rhodamine	Thermo Fisher Scientific Inc., Darmstadt, Germany
Dimethyl sulfoxide (DMSO)	Sigma Aldrich Chemie GmbH, Steinheim, Germany
Disodium phosphate (Na_2HPO_4)	Sigma Aldrich Chemie GmbH, Steinheim, Germany
Disodium phosphate (Na_2HPO_4)	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
ELISA blocking buffer 1 (concentration ELISA)	1x PBS, 0.05% (v/v) Tween®20, 1 mM EDTA
ELISA blocking buffer 2 (antigen ELISA)	1% BSA in PBS
ELISA blocking buffer 3 (serum ELISA)	1x PBS, 0.05% (v/v) Tween®20, 4% BSA
Elution buffer pH 3 (antibody purification)	0.1 M glycine pH 3.0
Neutralization buffer pH 9.0 (antibody purification)	1M Tris
Ethanol	Sigma Aldrich Chemie GmbH, Steinheim, Germany
Ethidium bromide	Sigma Aldrich Chemie GmbH, Steinheim, Germany

Ethylenediaminetetraacetic acid (EDTA)	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
GIBCO 10x PBS (pH 7.4)	Life Technologies GmbH, Karlsruhe, Germany
GIBCO 1x PBS (pH 7.4)	Life Technologies GmbH, Karlsruhe, Germany
GIBCO Trypan Blue Stain 0.4%	Life Technologies GmbH, Karlsruhe, Germany
Glycerol ($C_3H_8O_3$)	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Glycine ($C_2H_5NO_2$)	Sigma Aldrich Chemie GmbH, Steinheim, Germany
H_2O_2	Th. Geyer GmbH & Co. KG, Renningen, Germany
Paraformaldehyde (PFA)	Alfa Aesar, Thermo Fisher (Kandel) GmbH, Karlsruhe, Germany
Percoll	GE Healthcare Life Sciences, Freiburg, Germany
Polyethyleneimine (PEI)	Sigma Aldrich Chemie GmbH, Steinheim, Germany
Protein G Sepharose Fast Flow	GE Healthcare Life Sciences, Freiburg, Germany
SeaKem LE Agarose	Cambrex Inc., Rockland, ME, USA
Sodium chloride (NaCl)	Sigma Aldrich Chemie GmbH, Steinheim, Germany
Trizma base ($C_4H_11NO_3$)	Sigma Aldrich Chemie GmbH, Steinheim, Germany
Tween 20	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Cell lines and media	
HC-04	BEI Resources, NIAID, NIH, Manassas, VA, USA, MRA-975, contributed by J. S. Prachumsri
FreeStyle 293-F	Thermo Fisher Scientific Inc., Darmstadt, Germany
Ampicillin	Sigma Aldrich Chemie GmbH, Steinheim, Germany
GIBCO DMEM	Life Technologies GmbH, Karlsruhe, Germany
GIBCO HEPES	Life Technologies GmbH, Karlsruhe, Germany
GIBCO L-glutamine	Life Technologies GmbH, Karlsruhe, Germany
GIBCO RPMI	Life Technologies GmbH, Karlsruhe, Germany
GIBCO Trypsin EDTA 1x	Life Technologies GmbH, Karlsruhe, Germany
FreeStyle293 Expression media	Thermo Fisher Scientific Inc., Darmstadt, Germany
EX-CELL 293 Serum-Free Medium	Thermo Fisher Scientific Inc., Darmstadt, Germany
Parasites	
Pf NF54	Kind gift of Dr. Robert Sauerwein, Radboud University Medical Centre, NL
Kits	
AlexaFluor 647 Antibody Labeling Kit	Thermo Fisher Scientific Inc., Darmstadt, Germany
NucleoBond Xtra Midi / Maxi	Macherey-Nagel GmbH & Co. KG, Düren, Germany
NucleoSpin 96 PCR Clean-Up	Macherey-Nagel GmbH & Co. KG, Düren, Germany
NucleoSpin Gel and PCR Clean-up	Macherey-Nagel GmbH & Co. KG, Düren, Germany
NucleoSpin Plasmid Kit	Macherey-Nagel GmbH & Co. KG, Düren, Germany
Enzymes and Reagents	New England Biolabs GmbH, Frankfurt am Main, Germany
AgeI, BsiWI, SalI, XhoI	
Buffer 3.1 10x	New England Biolabs GmbH, Frankfurt am Main, Germany
Cutsmart buffer 10x	New England Biolabs GmbH, Frankfurt am Main, Germany
T4 DNA Ligase	New England Biolabs GmbH, Frankfurt am Main, Germany
T4 DNA Ligase buffer 10x	New England Biolabs GmbH, Frankfurt am Main, Germany
10x PCR buffer	Quiagen AG, Hilden, Germany
5x First strand buffer (RT)	Life Technologies GmbH, Karlsruhe, Germany
DTT	Life Technologies GmbH, Karlsruhe, Germany
Hotstart Taq DNA polymerase	Quiagen GmbH, Hilden, Germany
NP-40	Sigma Aldrich Chemie GmbH, Steinheim, Germany
Nuclease free water	Eppendorf AG, Hamburg, Germany
RNAsin	Promega Inc., Madison, WI, USA
RT buffer	Life Technologies GmbH, Karlsruhe, Germany
SuperScript III Reverse Transcriptase	Life Technologies GmbH, Karlsruhe, Germany
Taq DNA polymerase	Self-made
1kb Plus DNA marker	New England Biolabs GmbH, Frankfurt am Main, Germany
Desoxynucleotide Triphosphates (dNTPs)	Life Technologies GmbH, Karlsruhe, Germany
Oligonucleotides	MWG Biotech AG, Ebersberg, Germany
Random Hexamer Primers	Roche Diagnostics GmbH, Mannheim, Germany
Expression vectors	
1.5 ml reaction tubes	Sarstedt AG, Nümbrecht, Germany
2 ml reaction tubes	Sarstedt AG, Nümbrecht, Germany

13 ml tubes	Sarstedt AG, Nümbrecht, Germany
96-well Multiply-PCR plate	Sarstedt AG, Nümbrecht, Germany
96-well skirted twintech PCR plate	Eppendorf AG, Hamburg, Germany
Alpha Imager 1220	Alpha Innotech Corporation Inc., San Leandro, CA, USA
Aluminum foil seal	4titude, Surrey, UK
AxioObserver Z1 fluorescence microscope	ZEISS Microscopy, Oberkochen, Germany
BD Aria II	BD Biosciences GmbH, Heidelberg, Germany
BD LSR II	BD Biosciences GmbH, Heidelberg, Germany
BD Microlance 3 30G x 1/2"	BD Biosciences GmbH, Heidelberg, Germany
BioPhotometer	Eppendorf AG, Hamburg, Germany
Bio-Spin chromatography columns	Bio-Rad Inc., Hercules, CA, USA
Cell culture 48-well plate	TPP, Trasadingen, Switzerland
Cell culture 96-well plate	TPP, Trasadingen, Switzerland
Cell culture 96-well plate (transparent bottom)	TPP, Trasadingen, Switzerland
Cell culture dish (150 mm)	BD Biosciences GmbH, Heidelberg, Germany
CellStar sterile serological pipettes	Greiner Bio-One GmbH, Frickenhausen, Germany
Centrifuge 5180R (rotor A-4-81)	Eppendorf AG, Hamburg, Germany
Centrifuge 5417R (rotor F-45-30-11)	Eppendorf AG, Hamburg, Germany
CO_2 Incubator CB210	Binder GmbH, Tüttlingen, Germany
CryoTube Vials	Thermo Fisher Scientific Inc., Darmstadt, Germany
Domed 12-cap strips (PCR tube strips)	Bio-Rad Laboratories GmbH, München, Germany
Electrophoresis chamber D3 (horizontal)	Thermo Scientific Inc., Rochester, NY, USA
ELISA plates (96-well, flat bottom)	Costar Inc., Corning, Action, MA, USA
FluoNunc Plates	Costar Inc., Corning, Action, MA, USA
FrameStar 384	4titude, Surrey, UK
Heraeus B5042 (Bacteria incubator)	Kendro Laboratory Products, Weaverville, NC, USA
Inoculating loops/needles, polystyrene	VWR International Inc., Brigdgeport, NJ, USA
Leica DM2000 LED	Leica Microsystems GmbH Wetzlar, Germany
M1000Pro plate reader	Tecan, Crailsheim, Germany
Mastercycler ep Gradient S	Eppendorf AG, Hamburg, Germany
Mastercycler Pro 384	Eppendorf AG, Hamburg, Germany
Multipette plus	Eppendorf AG, Hamburg, Germany
Multitron Pro (Bacteria shaker)	Infors HT, Bottmingen, CH
Neubauer Counting Chamber by Marienfeld	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Omnican 50 Insulin Syringes, 12 mm 30 G	B. Braun Medical Ltd, Sheffield, UK
Petri dishes (100 mm)	Greiner Bio-One GmbH, Frickenhausen, Germany
Pipetboy acc	Integra Biosciences GmbH, Fernwald, Germany
Polypropylene tubes (15 ml, 50 ml)	Sarstedt AG, Nümbrecht, Germany
Polystyrene round bottom tube (5 ml) with cell strainer cap	BD Biosciences GmbH, Heidelberg, Germany
Slide-A-Lyzer Mini Dialysis Devices	Thermo Scientific, Rockford, IL, USA
S-Monovette EDTA K	Sarstedt, Nümbrecht, Germany
SpectraMax 190 Microplate Reader	Molecular Devices Inc., Sunnyvale, CA, USA
Stemi 2000 Stereomicroscope	ZEISS Microscopy, Oberkochen, Germany
Stuart Gyro rocker SSL3	Sigma Aldrich Chemie GmbH Steinheim, Germany
T75 cm^2 flask	BD Biosciences GmbH, Heidelberg, Germany
Thermomixer comfort	Eppendorf AG, Hamburg, Germany
Vortex genie 2	Scientific Industries Inc., Bohemia, NY, USA
Water bath with thermostat	JULABO Labortechnik GmbH, Seelbach, Germany
Wax seal	4titude, Surrey, UK
Software and webservers	
FlowJo v10.0	Treestar Systems Inc., Ashland, USA
GraphPad Prism 6.07	GraphPad Software Inc., La Jolla, USA
Illustrator CS5	Adobe Inc., San Jose, CA, USA
Pearl	Version 12.5, open source distribution
Python	Version 2.7.9, Python Software Foundation
Spyder	Version 2.3.5.2
Texmaker	Version 4.4.1
JabRef	JabRef version 2.10
Ensemble Genome Browser	release 91 [131]