Development and application of high-throughput screening methods to generate novel HIV-1 envelope immunogens

Vorgelegt von

Veronika Roswitha Graßmann

aus

Bruck in der Oberpfalz

im Jahr

2017

Development and application of high-throughput screening methods to generate novel HIV-1 envelope immunogens

DISSERTATION ZUR ERLANGUNG

DES DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT.)

DER NATURWISSENSCHAFTLICHEN FAKULTÄT III –

BIOLOGIE UND VORKLINISCHE MEDIZIN

DER UNIVERSITÄT REGENSBURG

Vorgelegt von

Veronika Roswitha Graßmann

aus

Bruck in der Oberpfalz

im Jahr

2017

Das Promotionsgesuch wurde eingereicht am:
03.07.2017

Die Arbeit wurde angeleitet von:
Prof. Dr. Ralf Wagner

Veronika Graßmann

Meinen Eltern

Table of Contents

1. Introduction

1.1. The HIV-1 pandemic

The human immunodeficiency virus type 1 (HIV-1) was first described as the causative agent of acquired immunodeficiency syndrome, short AIDS, more than thirty years ago.[1,2] Today, we know that four independent cross-species transmissions between apes and humans, most likely caused by bushmeat hunting, have led to the emergence of HIV-1, which can be subdivided into groups M (major), O (outlier), N (non-M/ non-O) and P (pending). While group M and N have been traced back to chimpanzee origin in Cameroon, the origin of group O is unknown and group P most likely derives from gorillas. Phylogenetic and statistical analyses have dated the first zoonotic transmissions to humans to the beginning of the 20th century. While the other groups have remained relatively small in comparison, HIV-1 group M since then has developed into a pandemic with millions of infections worldwide.[3] It is further subdivided into 9 different phylogenetic subtypes, the clades A to D, F to H, J and K, with A, B and C globally being the most prevalent.[4] While clade A viruses are predominant in eastern Europe and eastern Africa, clade B prevails in western and central Europe, America and Australia. Subtype C accounts for almost 50 % of HIV-1 infections worldwide and is the major clade in countries which harbor 80 % of all global infections, such as Africa and India. In addition, the number of circulating recombinant forms (CRFs), which today constitute an estimated 20 % of infections, is continuously increasing.[5]

According to the "Joint United Nations Program on HIV AIDS" (UNAIDS) and the WHO, about 35 million people have died of HIV since the beginning of the epidemic and approximately 37 million people are currently living with HIV. However, the worldwide distribution still varies considerably. Nearly 70 % of HIV infected individuals are living in Sub-Saharan Africa, accounting for an infection rate of approximately 1 in every 25 adults (**Figure 1**).[6,7]
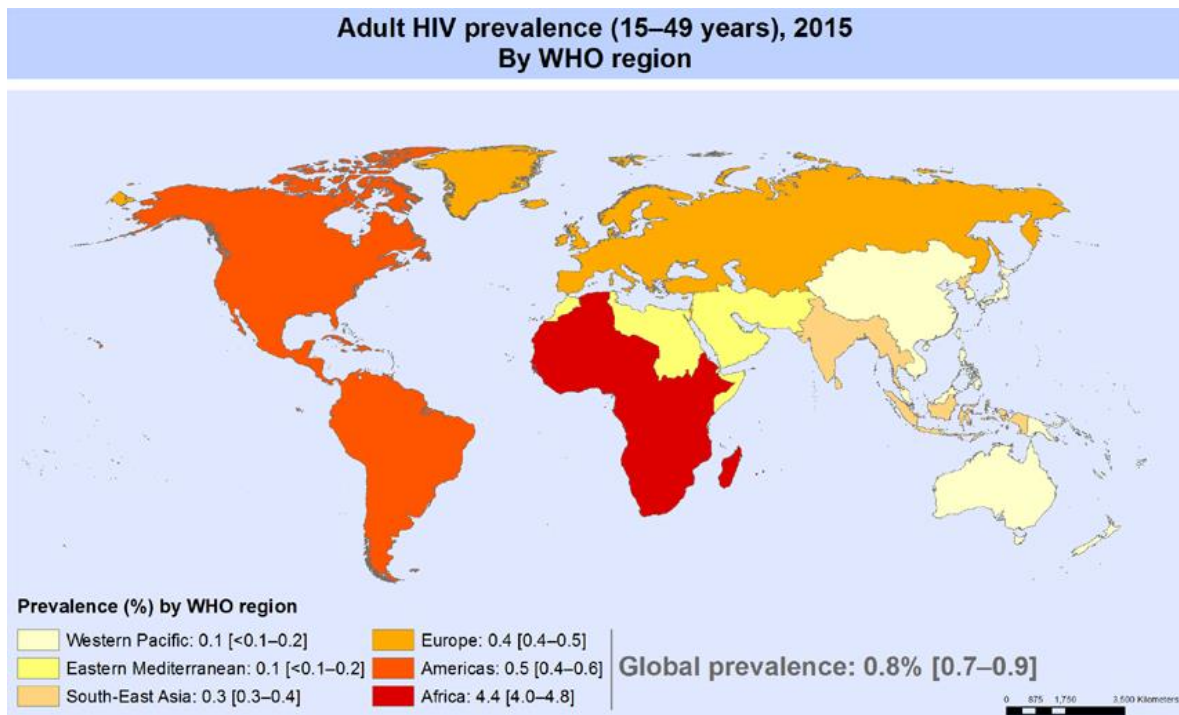
Figure 1: Global prevalence and distribution of HIV. More than 70 million people have already been infected with HIV, causing about 35 million AIDS related deaths. Globally, approximately 37 million people were living with HIV at the end of 2015, meaning that an estimated 0.8 % of adults aged 15 – 49 years were infected. The burden of the epidemic still varies considerably between regions, with Sub-Saharan Africa accounting for nearly 70 % of infections worldwide. From [6].

Although eradication of HIV-1 infection from patients to avoid the burden of lifelong therapy is the subject of extensive research (reviewed in [8]), a cure has only been reported once, in an acute myeloid leukemia patient receiving stem cell transplantation from a donor with a homozygous deletion in the CCR5 chemokine coreceptor used by HIV.[9] The difficulty for HIV-1 eradication in an individual is caused by the formation of HIV-1 reservoirs, which can develop within days of acute infection.[10] Although transcriptionally silent and therefore not visible for the immune system or anti-retroviral treatment, the virus remains replication-competent within these cells. Hence, if anti-retroviral treatment is interrupted, these reservoirs are able to reignite new rounds of infection.[8] The main reservoirs for HIV-1 represent latently infected resting memory CD4$^+$ T cells. However, other cell populations like latently infected naïve CD4$^+$ T cells, CD4$^+$ memory stem cells or macrophage and monocyte populations may contribute.[8] The potential reactivation of the virus in these cells poses the greatest obstacle in HIV-1 eradication, as each latently infected cell would have to be cleared to achieve a cure from HIV-1 infection.[8]

However, anti-retroviral therapy for HIV infected individuals has made significant progress during the last years. This, for example, accounts for the reduction of mother to child transmissions by 58 % from 2000 to 2014[11] and an almost normal life expectancy of infected individuals,[12] making HIV-1 a chronic infection for individuals with constant access to modern treatment. This, however, requires a high level of compliance with daily doses of pharmaceuticals.[13] More importantly, still only about 49 % of infected individuals have access to the treatment.[7] Especially for low income countries, this highlights the need for an efficient HIV-1 prophylactic vaccine in order to be able to eradicate the HIV-1 pandemic.

1.2.　The HI Virus: particle structure and replication cycle

HIV-1 is a retrovirus of the *Retroviridae* family with a diameter of about 100 to 120 nm. Viruses are enclosed by a lipid bilayer derived from the human cytoplasmic membrane. The group specific antigen (gag) proteins constitute the structural proteins of the viral particles. Myristinylated matrix protein gag p17 forms a layer on the inner side of the membrane, while gag protein p24 is arranged into the conical capsid (core). The capsid contains the two 9.2 kb RNA (+) strands that make up the viral genome, which are complexed with nucleocapsid protein gag p7. In addition, the core contains the viral enzymes (pol proteins) reverse transcriptase, integrase and protease. HIV-1 virus particles only possess one virus derived surface protein, the Envelope or short Env protein (**Figure 2**).[14]
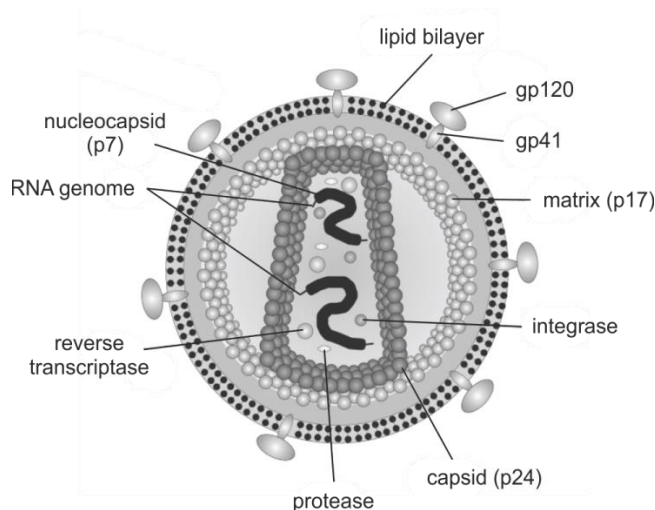
Figure 2: Schematic structure of a mature HIV-1 particle. HIV-1 is an enveloped virus with one viral surface protein, the envelope (Env), a trimer of three gp120 and three gp41 subunits. Each viral particle contains about 14 to 20 Env spikes. The inner layer of the host derived membrane is lined with a matrix of p17. The conical capsid of p24 protein contains the viral genome, consisting of two copies of (+) strand RNA complexed with nucleocapsid protein p7. HIV-1 particles also contain the viral enzymes reverse transcriptase, integrase and protease, as well as several other viral accessory proteins such as Vif, Vpr or Nef (not shown). Adapted from [180].

Each viral particle contains about 14 to 20 envelope spikes consisting of three gp41 and three gp120 subunits. Env initiates host cell infection by binding to the CD4 receptor on the surface of CD4$^+$ immune cells. This leads to a conformational change in Env which exposes its coreceptor binding site (**Figure 3**; step 1). Possible coreceptors for HIV-1 are the chemokine receptors CCR5 or CXCR4. Coreceptor binding initiates fusion of the viral and host membranes and ultimately delivers the viral capsid into the cytoplasm (**Figure 3**, step 2). After capsid uncoating (**Figure 3**; step 3), the reverse transcriptase transcribes the viral RNA genome to double-stranded DNA (dsDNA) (**Figure 3**; step 4). Subsequently, the pre-integration complex (PIC) consisting of dsDNA and associated proteins (**Figure 3**; step 5) is transported into the nucleus (**Figure 3**; step 6). Here, the viral integrase inserts the dsDNA into the euchromatin of the host genome (**Figure 3**; step 7) and thus triggers the transcription of viral genes by the cellular RNA polymerase II (**Figure 3**; step 8). During the early phase of transcription, multiply spliced mRNAs are generated, encoding the regulatory proteins Tat (trans-activator of transcription) and Rev (regulator of expression of virion proteins) (**Figure 3**; step 9). After re-entering the nucleus (**Figure 3**; step 10), Tat further activates viral transcription, while Rev participates in the export of singly spliced and unspliced transcripts to the cytoplasm (**Figure 3**; step 11). During the late phase of replication, these mRNAs produce the remaining viral proteins (**Figure 3**; step 12). While the glycoprotein Env is translated into the ER and accumulates in the plasma membrane via the Golgi export pathway, all other viral proteins are translated in the cytoplasm. Group specific antigens (gag) and viral enzymes (pol) are translated as gag and gag/pol precursor proteins. After myristinylation, those are transported to the plasma

membrane, where the assembly of the virions takes place (**Figure 3**; step 13). Via a psi (Ψ) sequence that only non-spliced, full-length RNAs contain, the viral genome binds to the gag precursor proteins. After budding of immature virus particles (**Figure 3**; step 14), the slightly acidic pH inside the particles leads to the activation of the viral protease, which in turn initiates virus maturation by cleavage of the gag and gag/pol precursors into the single proteins (**Figure 3**; step 15).[14,15]
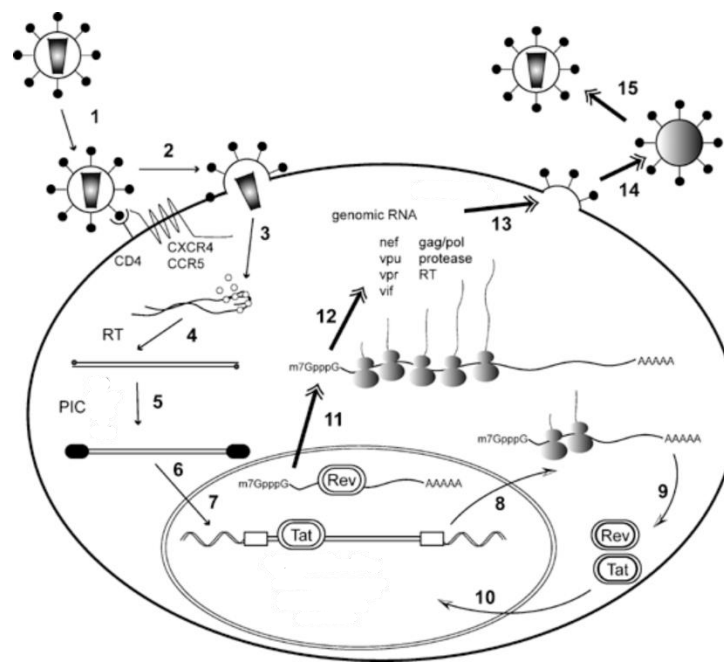
Figure 3: Schematic representation of the HIV-1 replication cycle. Small arrowheads: viral entry to integration. Curved arrows: early replication. Double-headed arrows: late replication. Binding of viral particles to the main receptor CD4 and the coreceptor (CCR5 or CXCR4) (1) leads to the fusion of viral and host membranes (2). After uncoating of the viral capsid (3), the viral reverse transcriptase transcribes the RNA genome into dsDNA (4). The pre-integration complex (PIC) consisting of dsDNA and associated proteins (5) is then transported into the cell nucleus (6). After integration of the viral dsDNA into the host euchromatin (7), transcription of viral factors starts with the early multiply spliced mRNAs (8). Translation of early regulatory proteins Tat and Rev (9), which are subsequently transported back into the nucleus (10), leads to increased transcription of viral mRNAs (Tat) and the export of singly spliced and unspliced viral mRNAs (Rev) (11). This results in the translation of all other virus encoded determinants, e.g. viral structural proteins (12). The assembly of viral particles takes place at the host plasma membrane (13). After budding of immature particles (14), a slightly acidic pH inside the viruses leads to the activation of the viral protease and subsequent maturation of the particles (15). From [15].

1.3. The HIV-1 envelope protein

1.3.1. Structure of HIV-1 envelope

The HIV-1 surface protein envelope (Env) is the only virus-encoded determinant present on the viral surface and thus visible to antibodies.[16] Each viral particle displays about 14 to 20 envelopes, which are synthesized as a gp160 precursor protein.[14] During its synthesis in the ER, gp160 is heavily glycosylated, with mainly high mannose and complex N-linked glycans finally accounting for about half of the molecular protein mass.[17] Golgi-associated furin proteases cleave gp160 into the two subunits gp120 and gp41. The functionally active envelope spikes are trimers of gp120-gp41 heterodimers with the three gp41 subunits forming the transmembrane-spanning region and the three gp120 subunits non-covalently attached on top.[18]

With an estimated 35 % sequence variation between the different subtypes, Env constitutes the most variable protein in the HIV-1 genome.[19] The gp120 subunit has a highly variable surface including five variable loops (V1–V5) (**Figure 4A** to **C** and **E**).[18] Structurally, gp120 can be separated into an inner and an outer domain, connected by the bridging sheet, a four-stranded β-sheet domain (**Figure 4D**). While the inner domain forms the non-covalent connection to gp41, the outer domain contains the CD4 receptor binding site.[20] Compared to gp120, gp41 is more conserved and consists of the fusion peptide (FP) and fusion peptide proximal region (FPPR), two heptad repeats (HR1 and HR2), the immunodominant C-C loop containing a conserved disulfide bridge, the membrane-proximal external region (MPER), the transmembrane domain (TM) as well as the cytoplasmic tail (CT) (**Figure 4A** to **C**).[18]
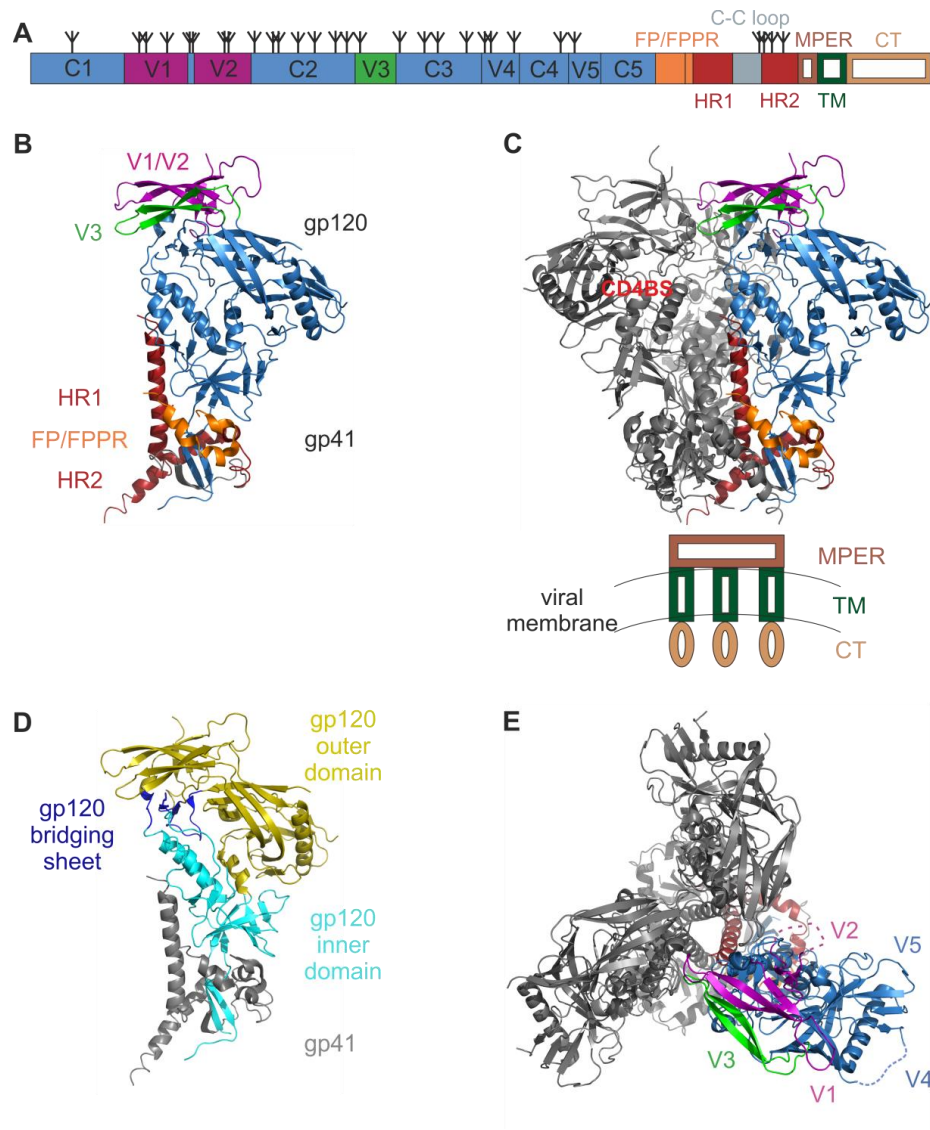
Figure 4: Structure of the HIV-1 envelope protein. Structures are based on BG505 SOSIP.664, PDB ID 4ZMJ. **A:** Schematic representation of HIV-1 envelope. Env is synthesized as a gp160 precursor protein and proteolytically cleaved into gp120 and gp41 by the Golgi-associated furin protease. gp120 consists of five constant regions (C1-C5), interspersed with five variable loops (V1-V5). gp41 contains the fusion peptide and fusion peptide proximal region (FP and FPPR), two heptad repeats (HR1 and HR2), the C-C loop with a conserved disulfide bond, the membrane-proximal external region (MPER), the transmembrane domain (TM) and a cytoplasmic tail (CT). Glycans are represented by tree-like symbols. **B:** Structure of one gp140 protomer. The gp41 heptad repeats 1 and 2 (HR1, HR2) are positioned at the base of the trimer. The fusion peptide proximal region (FPPR) and fusion peptide (FP) are located at the gp120 interface and the variable loops V1, V2 and V3 are positioned at the top of the protomer. **C:** Side view of the Env trimer. CD4BS: Location of the CD4 binding site on one protomer. Structures of the membrane-proximal external region (MPER), transmembrane domain (TM), and cytoplasmic tail (CT) are schematically included, as they are not solved in the crystal structure. **D:** Structure of a gp140 protomer subdivided into the outer and inner domain, which are connected by the four-stranded bridging sheet. **E:** Top view of the Env trimer. The variable loops (V1–V5) are located at the apex of the trimer and project outward. V2 and V4 are indicated with broken lines as they are disordered in the structure. Freely reproduced from [18].

1.3.2. Function of HIV-1 envelope

HIV-1 Env is the mediator of host cell entry. The first essential step of virus-host interaction constitutes the binding of the CD4 binding site on gp120 to the primary receptor CD4, a member of the immunoglobulin superfamily assisting the T-cell receptor in communicating with an antigen presenting cell.[21,20] This makes CD4$^+$ T cells as well as monocyte/macrophage lineages the target cells for HIV-1. Binding of Env to CD4 leads to a conformational rearrangement of the pre-fusion closed envelope spike. This includes the separation of V1/V2 from V3, resulting in the orientation of V3 towards the target cell membrane, as well as a rearrangement of the bridging sheet.[22] These conformational changes expose the coreceptor binding site on Env, which then interacts with one of two possible coreceptors, the chemokine receptors CCR5 or CXCR4. This interaction brings viral and host membranes into close proximity and induces exposure of the hydrophobic gp41 fusion peptide. Insertion of the fusion peptides of each subunit into the host membrane tethers virus and host. Subsequently, folding of the fusion peptides at a hinge region brings the amino-terminal helical region HR1 and the carboxy-terminal helical region HR2 together, resulting in the arrangement of a 6-helix bundle (6HB).[22] The 6HB is the driving force that brings both membranes closely together, allowing the formation of the fusion pore (**Figure 5**). In summary, coreceptor binding unlocks the potential energy of the gp41 fusion complex, resulting in 6HB formation, opening and stabilization of the membrane fusion pore, and subsequent delivery of the viral contents into the host cell cytoplasm.[21]
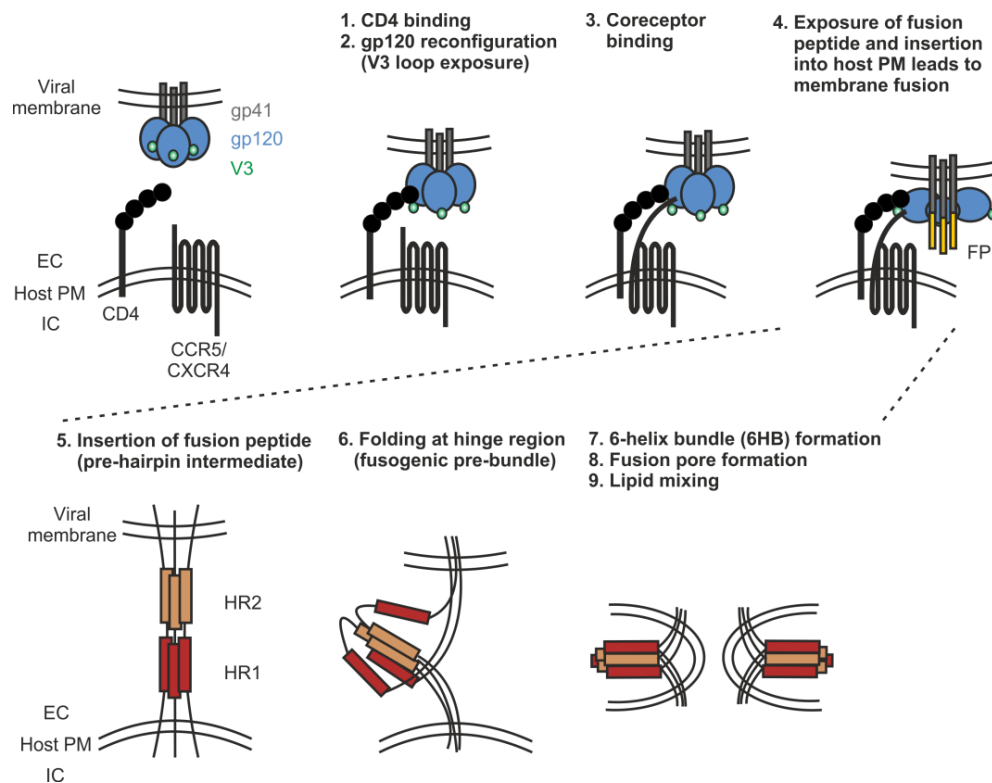
Figure 5: Schematic representation of HIV-1 cell entry. A functional, trimeric envelope spike initiates host cell infection by binding to the CD4 receptor on CD4⁺ cells (1). CD4 binding results in a conformational rearrangement of Env, including the exposure of V3 and the coreceptor binding sites (2). Suitable coreceptors are the chemokine receptors CCR5 or CXCR4. Coreceptor binding (3) leads to the exposure of the fusion peptide (4), which is then inserted into the host cell membrane (5). The fusion peptide of each gp41 subunit then folds at a hinge region (6), bringing heptad repeats HR1 and HR2 together to form a 6-helix bundle (6HB) (7). As the HR1 domains are in close proximity to the host cell membrane, and the HR2 domains are in close proximity to the viral membrane, the 6HB is the driving force that brings the opposing membranes together, resulting in the formation of a fusion pore (8) and subsequent delivery of the viral contents into the host cell cytoplasm (9). Grey: gp41, blue: gp120, green: V3, yellow: fusion peptide (FP), deep red: heptad repeat 1 (HR1), brown: heptad repeat 2 (HR2). EC: extracellular space, IC: intracellular space, host PM: host plasma membrane. Freely reproduced from [23].

1.3.3. Immune evasion mechanisms of HIV-1 envelope

The envelope protein is the only virus encoded antigen on the surface of infectious HIV-1 particles[14] and is therefore exposed to strong selection pressure by the host immune system.[24,25] As such, it has developed several sophisticated mechanisms to escape an efficient humoral immune response:

(i) Env variability (**Figure 6B**)**.** With about 35 % sequence variation between the different subtypes, Env constitutes the most variable protein in the HIV-1 genome.[19] It has been

estimated that the total hemagglutinin (HA) sequence variability for one year of an influenza epidemic is equivalent to the degree of HIV-1 Env sequence variability in a single individual.[19] The five variable loops of gp120 are able to tolerate high mutation rates without altering viral fitness and possess a high level of flexibility. Under immune pressure, they are able to change length, amino acid composition as well as the number of glycosylation sites to quickly escape neutralizing antibody responses.[18,26,27] The length of V1/V2, for example, can range from 50 to 90 amino acids (aa), while loops V4 and V5 differ from 19 to 44 aa and from 14 to 36 aa, respectively.[28,27]

(ii) Glycan shield (**Figure 6A**)**.** The highly glycosylated surface of gp120 also serves as a physical barrier to the immune system. Although many glycosylation sites are highly conserved, the exact number of glycans varies among the different isolates. In general, about 50 % of the Env molecular mass is made up of N-linked glycans, mostly of the high mannose and complex type. The resulting glycan shield covering the surface of Env limits the accessibility of conserved sites. In addition, the on-demand repositioning of glycans obstructs Env recognition by neutralizing antibodies.[29]
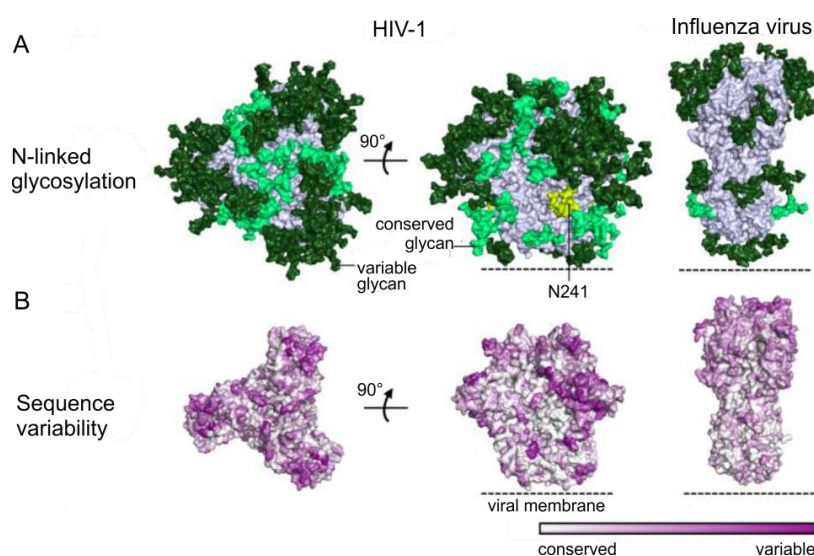
Figure 6: Glycan shield and sequence variability of HIV-1 envelope compared to Influenza hemagglutinin. **A**: N-linked glycosylation. Conservation: light green >90 %; dark green <90 %; yellow: conserved glycan at N241 not present in BG505. **B**: Sequence variability. Structures are based on BG505 SOSIP.664 (PDB 4TPV) and Influenza hemagglutinin H3 (PDB 2YP7). With minor modifications from [22] with permission of the Nature Publishing Group.

(iii) Conformational masking. The variable loops lie predominantly on the surface of gp120 (**Figure 4B, C** and **E**).[18] All conserved regions of gp120 are located underneath the V1/V2/V3 cap or near the interface of the subunits, as e.g. in the case of the CD4 binding site. This limits the accessibility for antibodies, allowing them to approach only at a defined angle. Glycans, e.g. the conserved glycan at N276, further restrict access to

conserved regions like the CD4 binding site.[18,30,31] In addition, the required two-step binding process to CD4 and the coreceptor, which triggers the spike-opening process of Env, is proposed to be an evolutionary adapted mechanism to protect conserved regions like the coreceptor binding site or the fusion machinery, which are prone to be targeted by neutralizing antibodies.[32] This way, they only become accessible when the virus is already in close proximity to the target cell and the fusion process is already under way.

(iv) Conformational flexibility. Env`s flexibility necessary for the profound conformational changes during host cell entry also contributes to immune evasion due to the multitude of Env conformations presented to the immune system (**Figure 7**). This is enhanced by the structural instability of Env. The non-covalent linkage of the gp120 subunits to the gp41 transmembrane stem can lead to shedding of gp120. Open Env trimers and shedded gp120 result in the exposure of immunodominant epitopes usually hidden in the inner part of the functional, closed trimer spike, which leads to the induction of various non-neutralizing antibodies that further hinder an efficient humoral immune response.[22]
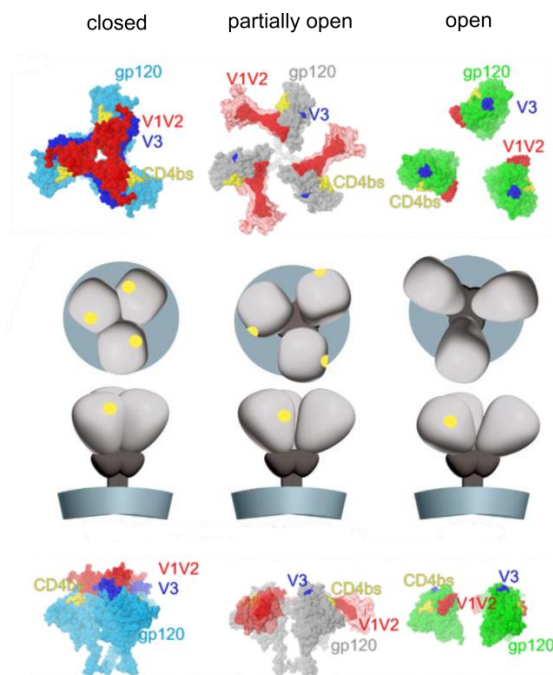
Figure 7: Conformational flexibility of HIV-1 Env. The inherent flexibility of HIV-1 Env necessary for the profound conformational changes during host cell infection as well as gp120 shedding present the immune system with a multitude of different Env conformations. Immunodominant epitopes usually hidden in the inner part of the functional, closed Env trimer are exposed, leading to the induction of various non-neutralizing antibodies. Yellow: CD4 binding site, middle panel: light gray: gp120, dark gray: gp41, blue-gray: membrane. Upper and lower panel: light blue: closed state; gray: partially open state; green: open state; red: V1/V2; dark blue: V3. middle panel: with permission from [110], upper and lower panel: from [181].

Taken together, these escape mechanisms account for the lack of an efficient humoral immune response against HIV-1. Although most infected individuals develop neutralizing antibodies during the first year of infection, those are mostly strain-specific and can

quickly be circumvented by escape mutations or other quasi-species present in an infected individual.[33] However, some infected individuals develop so called broadly neutralizing antibodies capable of neutralizing a broad panel of different isolates.[33] The different types of antibody responses during HIV-1 infection are described in more detail in the following section.

1.4. Antibody responses against HIV-1

1.4.1. Timelines and stages of the antibody response against HIV-1

The antibody response against HIV-1 can be divided into three major groups as reviewed in [34]: Env-binding but non-neutralizing antibodies can be detected as early as one week after infection. Generally, antibodies against gp41 arise first, followed by antibodies targeting variable regions on gp120, e.g. V3. However, these antibodies do not seem to put selection pressure on the virus, nor affect plasma viral load[35,33] and hence might mostly target dissociated gp120 and gp41 proteins or aberrantly folded trimers.[36] Although non-neutralizing, it is likely that some of these antibodies can convey Fc-mediated effector activities like ADCC (antibody-dependent cellular cytotoxicity) or ADCP (antibody-dependent cellular phagocytosis) and thus they are also of functional interest.[34]

Several weeks or months after HIV-1 infection, autologous neutralizing antibodies (nAbs) can be detected. The reason for their slow development is still unclear, but might be associated with impaired $CD4^+$ T cell support in antibody maturation due to the HIV-1 infection or the various immune escape mechanisms of HIV-1 Env (see 1.3.3). The very narrow neutralization profile, which is mostly limited to the autologous virus, only leads to the selection of other quasi-species not targeted by those nAbs or the insertion of escape mutations to avoid the neutralizing antibody response.[34]

However, an ongoing co-evolution of escaping viruses and the humoral immune system adapting to these escape variants ultimately leads to increased neutralization breadth, resulting in broadly neutralizing antibodies (bnAbs) in about 10 - 50 % of infected individuals after several years of infection.[37,38] The most potent bnAbs known today can neutralize more than 90 % of the different isolates tested.[39]

1.4.2. Broadly neutralizing antibodies

Broadly neutralizing antibodies can neutralize a variety of different HIV-1 isolates, with the best among them being able to neutralize more than 90 % of the tested isolates[39]. In comparison to other antibodies, bnAbs share several unique features (reviewed in [39,38]):

(i) High amount of somatic hypermutation: Due to their long evolution time of several years, a commonly shared feature of HIV-1 bnAbs is their high level of somatic hypermutation (SHM). While most affinity-mutated V_H domains carry between 15 and 20 mutations at nucleotide level, HIV-1 bnAbs can exhibit 40 to 100 V_H gene mutations.[32]

(ii) Long CDR H3 loops: Several bnAbs also display long variable heavy-chain third complementary-determining regions (CDR H3). While the median length of CDR H3 lies at 13.5 amino acids, many HIV-1 bnAb CDR H3 loops reach up to 20 to 34 residues,[40] often enabling these antibodies to penetrate the glycan shield of Env.[32] It is suggested that the development of such antibodies requires the engagement of rare germline B cells.[38]

(iii) Poly- and autoreactivity: Third, bnAbs often display a certain level of poly- and/or autoreactivity.

Taken together, these unusual features may pose some of the reasons why bnAbs only occur in approximately 10 to 50 % of infected individuals,[37,38] as exceptionally high levels of SHM, long CDR H3 loops, as well as poly- and autoreactivity are associated with negative B cell selection, limiting the pool of B cells able to develop bnAbs (reviewed in [32,40,41]).

Before 2009, only a handful of bnAbs were known, all targeting three major antigenic sites: (1) the CD4 binding site (e.g. VRC01[42], PGV04[43] or HJ16[44]), (2) a glycan epitope in the outer domain of Env (2G12[45]), and (3) the membrane-proximal external region (MPER) in gp41 (e.g. 2F5, 4E10[46,47] or 10E8[48,]). In recent years, however, the number of known bnAbs has increased drastically for several reasons. Due to the complex nature of HIV-1 Env (see 1.3), only core gp120 Env variants missing the variable loops were available for the analysis of sera for a long time. Today, new immunogens without any deletions and depicting native-like conformations are available.[49,50] In addition, high-throughput screening methods were developed to select broadly neutralizing B cells from the B cell repertoire of infected patients, e.g. using the newly available immunogens as baits in antigen-specific B-cell sortings[51,52,53] or via high-throughput neutralization assays of immortalized B cell repertoires.[54,55] This led to the discovery of a variety of new bnAbs with target sites distributed over the whole Env surface (**Figure 8**). One of the newly

identified antigenic sites includes a quaternary structure-dependent epitope spanning V1/V2 and specific N-linked glycans, which is e.g. targeted by bnAbs PG9, PG16 or PGT145.[56,57,58] Other new target sites include V3 with (PGT121-134)[55] or without (HGN194)[44] recognition of the conserved glycan at position 332 and glycan-dependent epitopes on the gp41-gp120 interface targeted by PGT151[59] or 35O22[60].
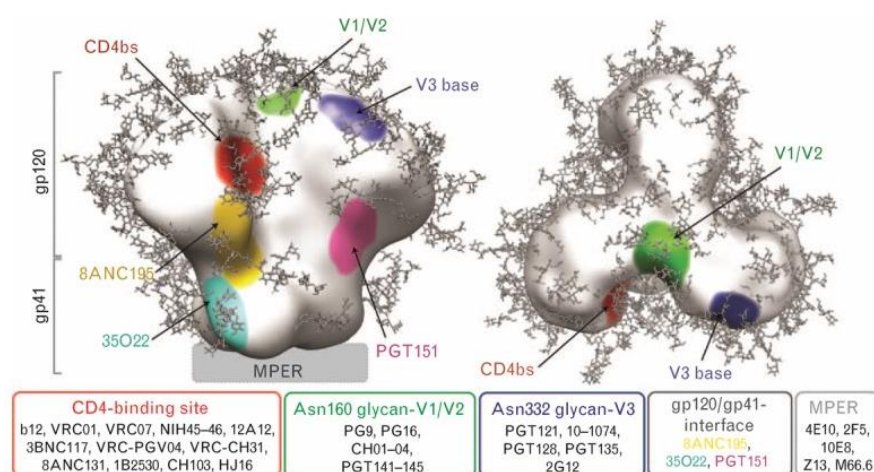
Figure 8: bnAb epitopes on HIV-1 Env. bnAb epitopes are distributed over the whole surface of Env and include the CD4 binding site, the trimer apex (V1/V2, including a conserved glycan at position N160), a N332 glycan-dependent V3 epitope, the gp120-gp41 interface, as well as the membrane-proximal external region of gp41 (MPER). From [61].

Remarkably, after passive immunization of macaques, these broadly neutralizing antibodies were shown to provide protection against chimeric simian/human immunodeficiency virus (SHIV) infection.[62,63] This led to the hypothesis that a preventive vaccine able to induce bnAbs would also result in protection against HIV-1 infection. However, none of the immunogens available to date were able to induce broadly neutralizing antibody responses *in vivo*. Hence, the currently most in-depth researched aspect of HIV-1 vaccine design is the development of HIV-1 immunogens and immunization strategies able to overcome this obstacle.

1.5.   HIV-1 envelope as an immunogen

1.5.1.   HIV-1 envelope laboratory adaptations

For analyses of HIV-1 Env, several truncated laboratory adaptations of the naturally occurring gp160 protein have been used over the years: for membrane-bound gp145, the cytoplasmic tail was removed to achieve enhanced expression and membrane incorporation.[64,65] To obtain soluble protein, either gp120 was employed[20,42] or Env was truncated N-terminally of the transmembrane domain to obtain soluble trimeric gp140 variants[66,49] (**Figure 9**). To avoid gp120-gp41 dissociation, the REKR furin cleavage site between gp120 and gp41 was often mutated to REKS or SEKS to obtain cleavage-incompetent envelope proteins.[67,68,69]
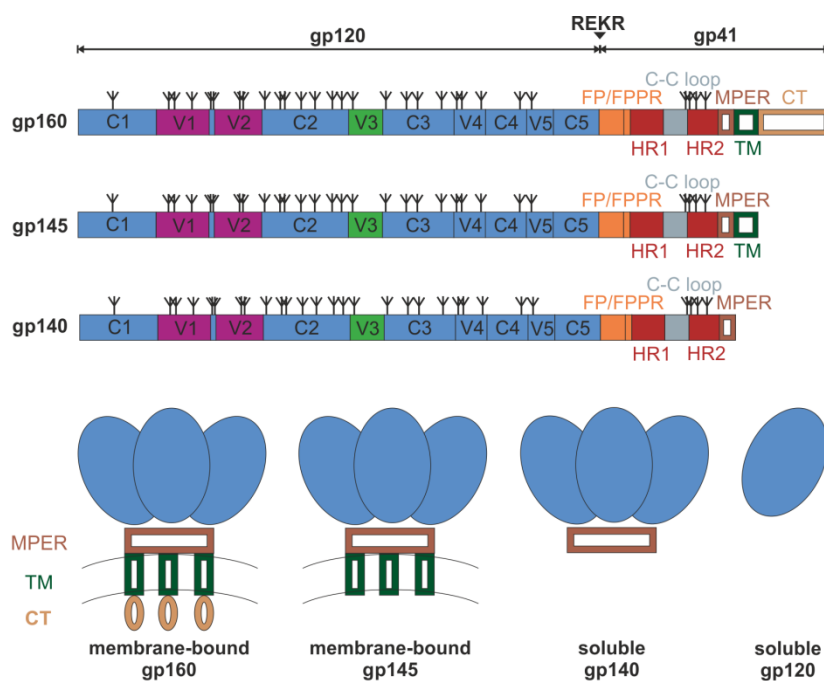
Figure 9: Laboratory adaptations of HIV-1 envelope. Truncation of the cytoplasmic tail (CT), resulting in membrane-bound gp145, led to increased Env expression and surface presentation. To obtain soluble Env gp140 trimers, the protein can be truncated N-terminally of the trans-membrane domain (TM). In addition, termination at the furin cleavage site (REKR) results in soluble gp120. (See Figure 4 for all other abbreviations in this schematic).

1.5.2. Clinical efficacy trials conducted until today

So far, six clinical efficacy studies have been completed (reviewed in [70] and [71]). The first efficacy studies for HIV-1 immunogens were started in the late 1990s with recombinant gp120 vaccines of bivalent subtype B/B for the North American trial (VAX 004) and bivalent subtype B/E for Thailand (VAX 003). The final results of these two trials indicated no protection against HIV infection.[72]

This outcome of the VaxGen trials was an incentive for intensified research on the cell mediated immune response, especially as cytotoxic T-lymphocyte (CTL) responses play an important role in controlling replication in infected people and hence early viral control. This was mostly pursued by developing DNA and viral vector immunizations. Efficacy trials were carried out using replication defective recombinant Adenovirus 5 (Ad5) vector vaccines with HIV-1 clade B gag/pol/nef inserts in the Step trial starting in 2004 and the Phambili trial starting in 2007. Although the vaccine used for the Step trial was found to produce reasonable CD8$^+$ T-cell responses, those were mainly directed against variable regions of the virus. In 2007, futility was acknowledged for the efficacy objective of both trials.

The Step trial showed that pre-existing immunity to Ad5 increased risk of HIV-1 acquisition. Therefore, the clinical trial HVTN505 evaluated the efficacy of a DNA prime - Ad5 vector boost vaccine in Ad5 seronegative individuals. Here, a plasmid mixture for clade B gag, pol and nef, as well as env from clades A, B and C was used, followed by an Ad5 vector boost to overcome the antibody and T-cell breadth problems demonstrated by the previous studies. However, this also did not result in efficacy against HIV infection or a reduction in viral load set point.

The negative results from those trials brought on a rethinking in the field to search for more strategic and systematic HIV-1 vaccine approaches. The RV144 or Thai trial, which was conducted from 2004 to 2009, produced the first positive, albeit modest, result for an HIV vaccine efficacy trial. RV144 tested a prime-boost combination of two vaccines: ALVAC, a canary pox vector from Sanofi Pasteur expressing clade E gp120 and clade B Gag and Pro, followed by an AIDSVAX B/E gp120 boost, which was already tested in the VaxGen trials without showing efficacy. In 2009, the trial demonstrated an efficacy of about 31 % in preventing infection after 3.5 years. The immune correlates analysis published in 2012 suggests that non-neutralizing antibodies directed against V1/V2 may contribute to protection by eliciting effector responses like ADCC (antibody-dependent

cellular cytotoxicity).[73,74] After these promising results, a pox-protein regimen adapted for Sub-Saharan Africa containing an ALVAC vector with clade C Env insert and a bivalent clade C recombinant gp120 will be tested in the HVTN702 trial starting in late 2016.[70]

1.5.3. Immunogen efforts for the induction of cross-neutralizing antibody responses

In order to develop a prophylactic vaccine based on the induction of broadly neutralizing antibodies, several strategies have been pursued and tested in animal models. They aim at overcoming the obstacles of an efficient neutralizing antibody response created by the various immune escape mechanisms of Env (see 1.3.3) and the resulting complex co-evolution necessary for the emergence of bnAbs *in vivo* (see 1.4).

Env variability and the resulting problem of antibody strain-specificity was e.g. addressed by a directed evolution approach screening a clade B chimeric gp120 Env library for improved binding to neutralizing monoclonal antibodies. Eventually, one of these variants, ST-008, was able to elicit improved neutralizing antibody responses in rabbits[75]. Other approaches used heterologous substitution or deletion of the V1 loop,[76] or hyperglycosylation of variable loops[77] to focus antibody responses to more conserved epitopes in soluble gp140 trimers. Although those approaches shifted immunogenicity in rabbits from the variable regions of Env to other epitopes like the CD4 binding site, the breadth of those antibody responses was limited, indicating the need for further modified Env variants.

To address the problem of Env`s structural flexibility, which results in the presentation of non-neutralizing epitopes usually hidden in the inner part of the trimer, gp140 trimers were chemically cross-linked.[78,79] Cross-linking largely conserved binding of bnAbs to their different epitopes, but reduced binding of several non- or weakly neutralizing antibodies and soluble CD4 (sCD4).[78] Also, cross-linking significantly improved Env thermostability and trimer integrity with or without adjuvant[50] and immunization studies in rabbits revealed significantly increased neutralization titers.[78,50]

The so-called SOSIP and NFL approaches for soluble trimers also address the problem of structural heterogeneity caused by Env flexibility. Previously used gp140 variants usually had a mutated furin cleavage site to prevent gp120 – gp41 dissociation.[80] Today, it is

known that this prohibits correct Env folding and leads to aberrantly folded Env trimers.[81] To overcome this problem, the cleavage-independent native flexibly linked (NFL) Env trimers were developed. With a glycine-serine (G4S) flexible peptide linker replacing the REKR furin cleavage site, the two subunits are able to fold homogeneously to form cleavage-independent, well-folded Env trimers. This is further supported by a trimer stabilizing proline mutation in gp41 (**Figure 10**). This strategy was successfully applied to several different HIV-1 subtypes.[82]

Similarly, the soluble recombinant clade A Env trimer BG505 SOSIP.664 was designed to form well-folded trimers. The SOSIP strategy implements an improved 6xR furin cleavage site instead of REKR, resulting in fully cleaved gp140 trimers. gp120 shedding is prevented by a novel disulfide-bridge between gp120 and gp41 (SOS) and trimerization is enhanced by a trimer stabilizing isoleucine to proline mutation in gp41 (IP) (**Figure 10**). Termination of gp140 SOSIP variants at amino acid 664 diminishes protein aggregation.[49]
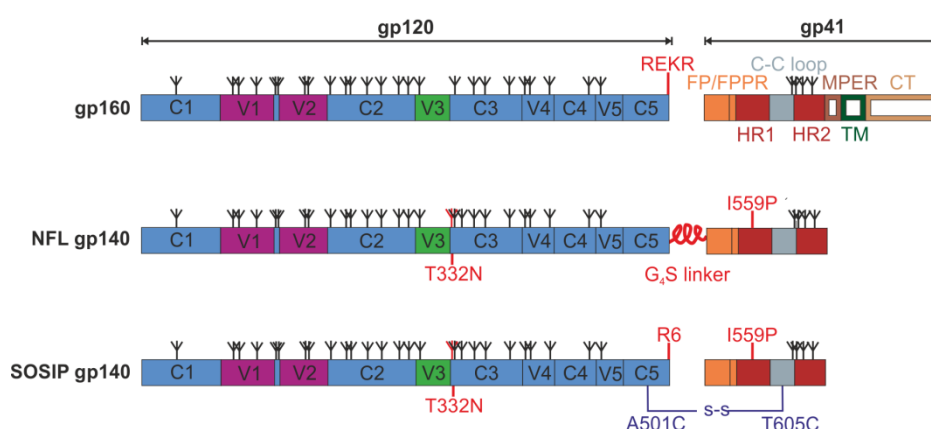
Figure 10: Design of NFL (native flexibly linked) and SOSIP Env variants. Mutation of the REKR cleavage site to prevent furin cleavage between gp120 and gp41 and hence subunit dissociation resulted in aberrantly folded Env trimers. Two different strategies were developed to solve this problem: In native flexibly linked (NFL) trimers, the REKR cleavage site is replaced by a glycine-serine (G4S) linker. Together with a trimer stabilizing isoleucine to proline mutation in gp41, this allows proper folding of the two subunits and results in cleavage-independent, homogeneously folded trimers. In the SOSIP approach, furin cleavage is even improved by replacing REKR with 6xR. This leads to complete cleavage between gp120 and gp41. Dissociation is prevented by a newly introduced disulfide-bridge between the two subunits (SOS) and trimerization is supported by a trimer stabilizing isoleucine to proline mutation in gp41 (IP). In isolates which do not naturally bear the N-linked glycan at position 332 necessary for V3/Glycan-dependent bnAbs, this glycosylation site was artificially included (T332N).

Whereas all other Env immunogens tested before were only able to induce neutralizing antibody responses against easy to neutralize tier 1 virus strains, BG505 SOSIP.664 was able to induce neutralizing antibodies against the more neutralization-resistant, sequence-matched tier 2 virus in mice, rabbits and macaques.[83,84] Hence, SOSIP variants currently represent the most promising approach on the way to bnAb induction and are further pursued in different ways: To broaden the neutralizing antibody response induced by one SOSIP variant, SOSIPs for several different clades have already been developed.[85,86] However, in comparison to the original BG505 SOSIP, which forms over 90 % well-folded trimers, SOSIP variants of other subtypes fold less well, requiring antibody-based positive or negative selection of well-folded trimers.[87] Hence, further efforts are made to identify additional stabilizing mutations. Insertion of an intra-protomer disulfide bridge that locks gp120 in the closed state as well as eight trimer-derived (TD) stabilizing mutations resulted in variants forming homogeneous and highly stable trimers no longer requiring positive or negative selection.[87] Simultaneous or sequential delivery of different clade A, B and C SOSIPs resulted in autologous tier 2 neutralization for each of the immunogens. However, heterologous tier 2 neutralization was only observed sporadically and with limited breadth.[88] Whether autologous nAbs can be the precursors of bnAb lineages is unclear, but insights into the natural progress of HIV-1 infection suggest that autologous nAbs may pressure the virus to expose immunotypes that trigger bnAb induction.[38,89] Hence, autologous tier 2 neutralizing titers may be a step towards a cross-neutralizing antibody response.

To mimic the complex co-evolution between Env and the immune system *in vivo*, sequential immunization with different Env immunogens was applied. These approaches intend to initially prime the germline receptor on B cells, and subsequently induce the mature bnAb over several intermediate stages.[90] A recent study reported a sequential immunization strategy aiming at VRC01-class CD4 binding site bnAbs. The engineered envelope variant eOD-GT8[91,92], which binds VRC01 germline antibodies due to eight germline-targeting mutations (GT8), was used as a priming immunogen in a transgenic mouse model expressing VRC01 germline-reverted heavy chain. Sequential immunization with a BG505 GT3 intermediate and a final boost with BG505 SOSIP N276D drove the maturation of primed B cells toward highly mutated VRC01-class nAbs.[93] Similarly, sequential immunization experiments for the glycan-V3-targeting PGT121 bnAb were reported. Immunization with a germline-targeting prime, followed by several boosting

immunogens with gradual changes in Env structure induced high levels of somatic hypermutation, which was directly related to the acquisition of neutralizing activity.[84]

Almost all of those strategies have one thing in common: the need for Env immunogens with high affinities towards the respective germline, intermediate or mature antibodies, representing the starting point of this thesis.

2. Objective

The lack of HIV-1 envelope immunogens able to induce broadly neutralizing antibody (bnAb) responses *in vivo* is still the major obstacle of HIV-1 vaccine development. Recent studies suggest that trimeric envelope proteins with improved stability and antigenicity (i.e. preserved or improved binding to broadly neutralizing antibodies and reduced binding to non-neutralizing antibodies) are able to induce enhanced neutralizing antibody responses.[83,93,50,78,69] However, regarding neutralization-resistant tier 2 viruses, which dominate human transmissions,[83] neutralization breadth is still mostly limited to the autologous virus strain.[83,88]

This thesis focused both on the development of a novel high-throughput screening approach for the identification of antigenicity-improved envelope vaccine candidates as well as on the rational design of next generation envelope immunogens. The first objective was to provide a proof of concept study to illustrate the capability of a mammalian cell display and cell sorting-based library panning approach to isolate trimeric envelope antigens according to their antibody binding properties. To this end, a model library of five chimeric Env variants was implemented. The second objective was to develop antigenicity and stability improved envelope immunogens. To this end, an Env alanine scanning approach in combination with mammalian cell display was to be applied to characterize Env – antibody interactions and identify mutations which improve bnAb recognition, while reducing binding of non-neutralizing antibodies. Subsequently, I aimed to further support antigenicity and stability of the resulting next generation envelope variants by chemical cross-linking and analyze the impact of these modifications on immunogenicity in a preclinical animal model.

3. Materials and Methods

3.1. Human codon optimization

All HIV-1 envelope variants in this thesis were optimized for human codon usage with the online tool "*Geneart Geneoptimizer*" (Thermo Fisher Scientific).

3.2. Molecular Biology

3.2.1. General Molecular Biology

Unless specified otherwise, cloning of all DNA constructs was performed using common molecular biology protocols.[94] For restriction digestion of vectors and inserts, endonucleases from New England Biolabs (NEB) or Fermentas (Thermo Fisher Scientific) were used according to manufacturer`s instructions.

Inserts were either generated by direct restriction digestion from the appropriate plasmid or a preparative PCR using Phusion High Fidelity Polymerase (NEB, # M0530). For the latter, the use of appropriate oligonucleotides (**Appendix Table 1**) ensured attachment of the respective endonuclease restriction sites needed for cloning and, where necessary, the integration of mutations via fusion PCR.

Amplified or restriction digested DNA was purified by $1 - 3$ % agarose gel electrophoresis (depending on the size of the DNA fragments) and subsequent purification of the respective bands using Qiaquick gel extraction kit (Qiagen, # 28706). For the ligation of restriction digested vectors and inserts, T4 ligase (NEB, # M0202) was used with a three-fold molar excess of insert in relation to the vector. For transformation, chemically competent or electro-competent *Escherichia coli* (*E. coli*) DH10B or DH5α were used. For vectors still containing a CcdB cassette[40], the CcdB resistant strain DB3.1 was applied. Analytical PCR to identify insert-containing colonies was carried out using GoTaq Green MasterMix (Promega, # M712).

The cloning into QuickLigation (QL) vectors containing a CcdB killer gene cassette was performed using golden gate cloning as described previously using a combined restriction and ligation process.[96] Shortly, a restriction mixture containing the QL plasmid, the respective DNA insert and the type 2S endonuclease Esp3I (Thermo Fisher Scientific, # ER0452) was incubated in a PCR cycler for 45 min at 37 °C. Subsequently, the ligation mixture was added and the reaction was further incubated at 25 °C, 37 °C and 25 °C for 45 min each.

Plasmids were isolated from LB cultures using GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific, # K0503) for Sanger sequencing purposes or Qiagen plasmid kits for higher DNA yield. For transfection of FreeStyle 293-F cells, DNA was purified using Qiagen Maxi or Mega Kits (# 12165 and 12183). DNA for transfection of HEK293T cells and subsequent FACS analysis was purified from 35 ml TB cultures using Qiagen Plasmid *Plus* Midi Kit (# 12945, high yield protocol). The purity and concentration of DNA was measured using a NanoDrop 1000 instrument (Thermo Scientific) and sequence correctness of all constructs was verified by Sanger sequencing (GATC, Konstanz, Germany; GeneArt, Regensburg, Germany or Seqlab, Göttingen, Germany). A detailed catalogue of all oligonucleotides used in this thesis is listed in **Appendix Table 1**, while vectors are shown in **Appendix Table 2**.

E. coli **DH5α**	supE44 ΔlacU169 (φ80 lacZ ΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1
E. coli **DH10B**	F– mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara leu) 7697 galU galK rpsL nupG λ–
E. coli **DB3.1**	F- gyrA462 endA1 glnV44 Δ(sr1-recA) mcrB mrr hsdS20(r$_B^-$, m$_B^-$) ara14 galK2 lacY1 proA2 rpsL20(Sm^r) xyl5 Δleu mtl1
LB (lysogeny broth) medium	1% NaCl; 1% tryptone; 0.5% yeast extract; pH 7.5
TB (terrific broth) medium	1.2% tryptone, 0.5% glycerol, 2.4% yeast extract, 0.17 M KH$_2$PO$_4$ and 0.72 M K$_2$HPO$_4$

3.2.2. Permutation of selected 96ZM651 positions

Sequential permutation of 96ZM651 Env gp145 at selected positions was performed by GeneArt, Regensburg, Germany. Variants were provided as pool of one position glycerol stocks. To separate the individual mutants, small amounts of the glycerol stocks were plated on LB$_{Amp}$ agar (100 µg/ml ampicillin) to obtain single clones. 96 single clones per permutated position were inoculated in 96-well LB$_{Amp}$ agar plates provided by GATC

Biotech, Konstanz, Germany and sent to GATC Biotech for 96-well Sanger sequencing (SupremeRun 96). After sequences and purified DNA were obtained back from GATC, one of each correct identified mutant was transformed back into *E.coli* DH10B and purified from 35 ml TB cultures using Qiagen Plasmid *Plus* Midi kit (# 12945, high yield protocol). Amino acid variants which could not be identified by this procedure were generated by overlap extension PCR using the reagents listed in 3.2.1 and appropriate oligonucleotides to insert the respective mutations (**Appendix Table 1**).

3.3.   Cell biology

3.3.1. Cultivation of human cells

The adherent eukaryotic cell lines HEK293T (ATCC, # CRL-11268), Flp-In™ T-REx™ 293 (Thermo Fisher Scientific, # R78007) or stable cell lines generated from Flp-In™ T-REx™ 293 cells were cultivated in Dulbecco's modified Eagle medium (DMEM, Gibco/Thermo Fisher Scientific, # 11995-065) at 37 °C and 5 % CO_2 according to common protocols with the respective supplements indicated in the table below. Cells were cultivated up to 80 % confluency and split at a ratio of 1:10 every two to three days by washing with phosphate buffered saline (PBS), detaching with Trypsin/EDTA solution (PAN-Biotech, # P10-023500) and resuspension in the respective culture medium. FreeStyle™ 293-F suspension cells (Thermo Fisher Scientific, # R79007) were cultivated at 37 °C, 8 % CO_2 and 90 rpm in a culture volume of 300 ml using unbaffled 1 l flasks (Corning, # CORN431147), at cell concentrations from 0.1 to 1.6 Mio/ml in the medium indicated below. For sub-culturing, the respective amount of cells was centrifuged at 100 g for 5 min at RT, before pellets were resuspended.

PBS	137 mM NaCl, 2.7 mM KCl, 10 mM Na$_2$HPO4, 1.47 mM KH$_2$PO4, pH 7.4
FCS for HEK293T cell cultivation	GIBCO, # 10270-106
Tetracycline free FCS for Flp-In™ T-REx™ 293 cell cultivation (FCSΔtet)	Biochrom, # S0115
Penicillin/Streptomycin (Pen/Strep)	10000 U/ml penicillin, 10 mg/ml streptomycin, PAN Biotech, # P06-07100
Culture Medium for HEK293T cells	DMEM + 10 % FCS + 1% Pen/Strep
Culture Medium for untransfected Flp-In™ T-REx™ 293 cell	DMEM + 5 % FCSΔtet + 1 % Pen/Strep + 100µg/ml zeocin + 15 µg/ml blasticidin
Culture Medium for stable cell lines generated from Flp-In™ T-REx™ 293 cells	DMEM + 5 % FCSΔtet + 1% Pen/Strep + 100µg/ml hygromycine + 15 µg/ml blasticidin
Culture medium for FreeStyle™ 293-F Cells	FreeStyle™ 293 Expression Medium (Thermo Fisher Scientific, # 12338018) + 0.5 % Pen/Strep

3.3.2. Transient transfection of adherent human cells

HEK293T cells were transiently transfected using polyethylenimine (PEI).[97] For 96-well transfection, $3 - 4 \times 10^5$ cells were seeded into a 96-well flat bottom plate the day before transfection to reach ca. 80 % confluency on the day of transfection. Directly before transfection, the medium was exchanged to 30 µl DMEM without supplements. After 10 min of incubation at RT, the transfection mix containing 200 ng DNA, 0.8 µl PEI (Polyethylenimine, linear, Polysciences, # 23966-2, 1 mg/ml) and 30 µl DMEM without supplements was added to the cells. Five to six hours post transfection, the medium was changed back to HEK293T culture medium. For 6-well transfection, 5×10^5 cells were seeded into each well the day before transfection. Directly before transfection, the medium was replaced with 1 ml DMEM without supplements. A transfection mixture containing 100 µl DMEM without supplements, 2 µg DNA and 10 µl PEI (Polyethylenimine, linear, Polysciences, # 23966-2, 1 mg/ml) was incubated for 10 min at RT and then added to the cells. Five to six hours post transfection, the medium was replaced with HEK293T culture medium. In the case of SOSIP gp145 variants, the applied DNA for transfection consisted of the respective Env-containing plasmid and pcDNA3.1-furin at a ratio of 3:1 (150 ng:50 ng for 96-well transfection and 1.5 µg:0.5 µg for 6-well transfection).

3.3.3. Flow cytometry analysis of transiently transfected cells

48 h after transient transfection, cells were washed with PBS and detached with FACS buffer (PBS + 1 % heat inactivated FCS + 2 mM EDTA). Cells were transferred to a 96-well round bottom plate and centrifuged at 500 g and 4 °C for 5 min. After removal of the supernatant, cells were stained for 1 h at 4 °C with the respective screening antibody (PG9-Alexa647: 5 µg/ml, all other Alexa647-labeled antibodies: 4 µg/ml) and reference antibody 5F3-PE (2.5 µg/ml) (antibodies and labeling kits listed in 3.4.6) in 30 µl FACS buffer. Subsequently, three washing cycles with 200 µl FACS buffer each followed (centrifugation at 500 g, 4 °C, 5 min). Soluble CD4 (sCD4) staining was carried out at a concentration of 1 µg/ml, followed by staining with 5F3-PE and the secondary antibody anti-human-CD4-APC (OKT4, eBiosciences, # 17-0048-42, 1:100 in FACS buffer), which has been shown to not interfere with the binding of sCD4 to gp120.[98] After the washing procedure, cells were analyzed using a FACS-Canto II (BD), Attune or Attune NxT device (Thermo Fisher Scientific). Untransfected cells were used as a negative control (n=1), while 96ZM651 wildtype transfected cells served as a reference (n=3) on each 96-well plate. Cells were gated for living, single cells displaying envelope expression, as indicated by 5F3 binding above envelope negative cells. The mean fluorescence intensity (MFI) ratio between the screening compound and the reference antibody 5F3 was calculated and set into relation to wildtype binding (WT = 100 %). 5 to 6 independent measurements were carried out per analysis.

3.3.4. 96ZM651 alanine scan

Prior to this thesis, the alanine mutagenesis library was obtained from GeneArt/Thermo Fisher Scientific, Regensburg, Germany using 96ZM651 gp145 (GenBank No: AF286224) as a template. The sequence was adapted to human codon usage, the cytoplasmic tail was truncated (aa 718-868) and a C-terminal HA-Tag (YPYDVPDYA) was added. In addition, the cleavage site was mutated from REKR to REKS. The template sequence (96ZM651 wildtype) was cloned into pWPXLd (Addgene plasmid # 12258) and alanine mutations were generated from amino acids 31 to 688 by fusion PCR, leaving cysteine residues unchanged. All mutants were verified by Sanger sequencing and plasmids were prepared

from 35 ml TB cultures using the high yield protocol of Qiagen Plasmid *Plus* Midi kit (#12945) according to manufacturer`s instructions. Concentrations of the alanine library were adjusted to 100 ng/µl in 96-well PCR plates. Transient transfection and FACS analysis of HEK293T cells was performed as described above (3.3.2 and 3.3.3) using a FACS-Canto II (BD). For statistical data analysis, R was used to apply a two-sided one sample t-test to the calculated percentage ratios against 1 for wildtype Env, while controlling the false discovery rate (FDR) to be smaller than 5 %, to adjust for the number of performed t-tests. Using PyMol (http://pymol.org/, download of educational version Sept. 14th 2010), significant gain or loss of binding mutations were plotted on the crystal structure of BG505 SOSIP.664 (PDB ID: 4ZMJ).

3.3.5. TZM-bl neutralization assay

Neutralization assays were carried out at the Duke Central Reference Laboratory, Durham, USA according to protocols for standardized assessments of neutralizing antibodies for HIV/AIDS vaccine development.[99] Serial dilutions of unlabeled and labeled bnAbs were tested in a TZM-bl assay against a selected panel of HIV-1 Env pseudoviruses.

3.3.6. Generation of stable cell lines using Flp-In™ T-REx™ 293 cells

Flp-In™ T-Rex™ 293 cells (Invitrogen/Thermo Fisher Scientific, # K6500-01) were used to generate stable cell lines with inducible envelope expression regulated under the Tet operator/repressor system. Cells were cultivated in DMEM supplemented with 10 % FCS (tetracycline free, Biochrom, # S0115), 1 % Pen/Strep (Pan Biotech, # P06-07100), 100 µg/ml zeocin and 15 µg/ml blasticidin. The day before transfection, $5x10^5$ cells were seeded into 6-well plates to reach 80 % confluency on the day of transfection. Directly before transfection, the medium was replaced with 1 ml DMEM without supplements. 2 µg of a plasmid mixture containing the respective pQL13-Env/V3 variant or the pQL13-Env/V3 library and the helper plasmid pOG44 carrying the integrase (1.6 µg pQL13-

Env/V3 variant or library + 0.4 µg pOG44) were mixed with 8 µl PEI (Polyethylenimine, linear, Polysciences, # 23966-2, 1 mg/ml) and 100 µl DMEM without supplements. The mixture was incubated for 10 min at RT and then given onto the cells. After 6 h of incubation, medium was changed to 2 ml DMEM supplemented with 5 % tetracycline free FCS and 1 % Pen/Strep. After 48 h, cells were harvested and seeded into a T75 flask. After additional 5 h, the medium was replaced by 10 ml DMEM supplemented with 10 % tetracycline free FCS, 1 % Pen/Strep, 100 µg/ml hygromycin and 15 µg/ml blasticidin to select cells with stable integration events. Selection was carried out for the following 25 days, with the medium being replaced every 72 h. Both stable cell lines from an equimolar plasmid mixture of all five variants as well as for each variant individually were generated.

3.3.7. TaqMan copy number assay

This assay was performed by Dr. Tim-Henrik Bruun and results are shown for better illustration of the panning procedure. Genomic DNA (gDNA) was prepared from all established cell lines (Qiagen, QIAamp DNA Blood Mini Kit, # 51104, blood and body fluids protocol) and samples were processed with TaqMan® Universal PCR Master Mix (Life Technologies/Thermo Fisher Scientific, # 4304437). The TaqMan® Copy Number Assay for Markers & Reporters (Life Technologies/Thermo Fisher Scientific, # 4400291) was used to determine the copy number of integrated pQL13 plasmids in all established stable cell lines. The TaqMan® Copy Number Reference Assay, TERT (telomerase reverse transcriptase), Human (Life Technologies/Thermo Fisher Scientific, # 4403316) was included to compare the number of insertions per cell to a genomic reference. Samples were analyzed in a StepOne Plus cycler (Applied Biosystems/Thermo Fisher Scientific) according to manufacturer`s instructions.

3.3.8. FACS analysis of the individual stable Env/V3 cell lines

Stable cell lines were induced to express GFP and Env by supplementing the medium with 1 µg/ml doxycycline (Dox.). After 24 h, cells were detached with FACS buffer (PBS + 1 % heat inactivated FCS + 2 mM EDTA) and centrifuged at 250 g for 5 min at 4 °C. Then cells were incubated with serial dilutions of Alexa Fluor 647 labeled (Thermo Fisher Scientific, Alexa Fluor 647 protein labeling kit, # A20173) 447-52D (Polymun, # AB014) or HGN194 (kindly provided by Prof. Dr. Antonio Lanzaveccia and Dr. Davide Corti) diluted in FACS buffer for 20 min at 4 °C to reach equilibrium binding. Subsequently, cells were washed three times with 1 ml ice cold FACS buffer (centrifugation for 5 min at 250 g and 4 °C), resuspended in 100 µl FACS buffer and subjected to cytometric analysis using a FACS Canto II (BD). As GFP and envelope expression were genetically coupled, the ratio of Env and GFP MFIs was calculated to normalize for varying Env expression levels (relative MFI). Negative controls (Env-negative cells titrated with the very same antibody concentrations) were substracted to set the starting point of the curves to 0, then the curves were fitted using non-linear least squares regression (hyperbolar one-site binding, GraphPad Prism 5.0). To analyze the ratio of Env and GFP expression in each stable cell line, cells were stained with Alexa Fluor 647 labeled mAb 5F3 (Polymun, # AB010, 5 µg/ml) as described above and MFI correlations of 5F3 to GPF were compared for each stable cell line.

3.3.9. FACS-panning: affinity enrichment of Env/V3 variants

30×10^6 cells of the Env/V3 cell library generated with an equimolar mixture of all five Env/V3 plasmids were seeded into a 15 cm dish and induced with doxycyline (1 µg/ml) the day before the panning procedure. After 24 hours, cells were stained with 1 ml Alexa Fluor 647 labeled 447-52D (2.7 nM) or HGN194 (6.7 nM) according to the protocol described in 3.3.8. Finally, cells were resuspended in FACS buffer (PBS + 1 % heat inactivated FCS + 2 mM EDTA) at a maximal concentration of 60×10^6 cells/ml and filtered with a 30 µm pre-separation filter (Miltenyi Biotech, # 130-041-407). 50 µl of sample were put aside to serve as an input control. Cells were sorted using a FACS Aria IIu (BD), with the instrument set to "single cell mode" to discard two-target-events and obtain the most

accurate counts for the sorting procedure.[96] Triangle shaped gates were applied to sort cells with the highest or lowest antibody affinity in relation to their GFP signal. At least 20,000 cells were collected per gate. The genomic DNA of both the sorted cells and the input control were prepared using QIAamp DNA Blood Mini Kit (Qiagen, # 51104, blood and body fluids protocol) according to manufacturer's instructions, but without addition of carrier DNA and elution in only 20 µl 10 mM Tris/HCl pH 8.0. The genomic DNA was then used as a template for nested-PCR (primers see **Appendix Table 1**) to amplify the integrated Env genes isolated in the sorting procedure for re-cloning into pQL13 and subsequent analysis by qPCR, capillary sequencing of single clones or next generation sequencing (NGS). In addition, the genomic DNA was also directly subjected to NGS.

3.3.10. Enrichment analysis using realtime PCR (qPCR)

To determine the distribution of Env variants in the re-cloned pQL13 plasmids after FACS sorting, a qPCR assay was used as described previously.[96] Shortly, the reverse-primers were designed to specifically bind one variant of the Env/V3 model library, whereas the probe and forward-primer were designed to bind every variant alike. Furthermore, an additional reverse-primer binding every variant alike was used as a reference primer, determining the total amount of applied Env genes. The reactions were performed according to manufacturer`s instructions (Thermo Fisher Scientific, DyNAmo Flash Probe qPCR Kit, # F-455). Primer sequences are listed in **Appendix Table 1**.

3.3.11. Enrichment analysis using next generation sequencing

NGS was performed on input and first round samples of the panning procedure using the Illumina MiSeq platform on both the envelope genes amplified from the gDNA and cloned back into pQL13, as well as directly on the genomic DNA of the selected cells and the input control. In both cases, Env/V3 loops were amplified in a two step PCR reaction to first add short universal adapter sequences on the 5´and 3´ends, and then add barcoded adaptor sequences according to TruSeq® Small RNA Sample Prep Kits (Illumina, Primer

sequences listed in **Appendix Table 1**). PCR products were purified using Agencourt AMPure XP magnetic beads (Beckman Coulter, # A63881) after each PCR reaction with a beads to sample ratio of 1 to 1.8. Sample preparation was performed in DNA LoBind tubes (Eppendorf, # 0030108051). AMPure beads were mixed with the DNA samples by pipetting thoroughly. After 5 min incubation at RT, samples were pulse-spinned and placed into a magnetic rack. After 3 min, supernatants were carefully removed and pellets were washed with 500 µl of freshly prepared 70 % ethanol. To this end, the tubes were turned around twice within one minute to move the beads around. After the suspensions cleared, supernatants were discarded. The washing procedure was repeated, then tubes were shortly centrifuged, placed back onto the magnetic rack and the remaining supernatant was carefully removed using a 20 µl pipette. The pellets were air-dried on the magnetic rack for 3 min. Subsequently, 20 µl of 10 mM Tris/HCl pH 8.0 were added and the pellets were dispersed by pipetting and vigorous mixing. Subsequently, tubes were pulse-spinned and placed back into the magnetic rack. After 3 min, the supernatants were transferred to new 1.5 ml LoBind tubes without disturbing the pellets. Quality control and quantification of PCR products was performed with a Bioanalyzer 2100 (Agilent) and High Sensitivity DNA Chips (Agilent, # 5067-4626). After quantification, samples were pooled and analyzed on a 300 cycles MiSeq V2 chip (Illumina, # MS-102-2002). To set up the sample sheet, the FastQ only, sample prep kit: true seq small RNA presetting was chosen. In the sample sheet, cycles were manually adjusted for 2x 151 reads to achieve sequencing from both the 5´ and the 3´ end. Sequence analysis was performed with CLC Main Workbench 7. After import into CLC, the files containing the associated forward and reverse reads were paired, then overlapping pairs were merged into one file without altering the predefined settings of the program. After export as FastQ files, distribution of the five Env/V3 chimeras within each sample was investigated with the V3 analyzer tool programmed by Dr. Benedikt Asbach.

3.4. Protein biochemistry

3.4.1. Construction of gp140 variants

gp140 constructs were obtained by amplification from the respective gp160 or gp145 containing plasmids. The transmembrane domain was omitted using suitable primers and, where necessary, the REKR cleavage site was mutated to REKS via overlap extension PCR (primers see **Appendix Table 1**). gp140 constructs were cloned into pQL13 or pcDNA3.1QL (**Appendix Table 2**) using golden gate cloning as described previously[96] and verified by Sanger sequencing (GeneArt, Regensburg, Germany or Seqlab, Göttingen, Germany).

3.4.2. Purification of soluble gp140 Env variants

Soluble gp140 proteins with or without His6-tag were purified from FreeStyle™ 293-F cells (Thermo Fisher Scientific, # R79007). Cells were cultivated according to manufacturer`s instructions in FreeStyle™ 293 expression medium (Thermo Fisher Scientific, # 12338018) supplemented with 0.5 % Pen/Strep (PAN Biotech, # P06-07100). For purification of soluble gp140 variants, 300 ml of HEK293-F suspension cells were grown to a cell concentration of 1 Mio/ml on the day of transfection. Before transfection, the $3x10^8$ cells were harvested at 100 g, 5 min and 4 °C and resuspended in 300 ml FreeStyle™ expression medium without supplements. 300 µg of Env gp140 containing plasmid DNA were mixed with 6.5 ml DMEM (Gibco/Thermo Fisher Scientific, # 11995-065) without supplements. For expression of SOSIP gp140 variants, 225 µg of Env SOSIP gp140 plasmid DNA and 75 µg of pcDNA3.1-furin were mixed. Simultaneously, 1.5 ml PEI (Polyethylenimine, linear, Polysciences, # 23966-2, 1 mg/ml) were diluted with 6.5 ml DMEM without supplements. Both mixtures were incubated separately for 5 min at RT and then mixed to be incubated for further 20 min at RT before the transfection mixture was added to the cells. After 6 h of incubation, the medium was changed to 300 ml of FreeStyle™ 293 expression medium supplemented with 0.5 % Pen/Strep. For higher protein yield, this procedure was scaled up using 3x 300 ml 293-F suspension cells. On day 5 after transfection, cells were harvested at 1000 g for 5 min at 4 °C and secreted gp140

variants were purified from the supernatant. To this end, the supernatant was filtered with a 0.22 µm filter (Sarstedt, Filtropur L 0.2 LS, # 83.1827.001) and loaded onto a 5 ml lectin affinity chromatography column (*Galanthus nivalis* (Snowdrop) lectin, Agarose bound, Biozol, # VEC-AL 1243-5) at 4 °C overnight with a maximal flow rate of 1 ml/min. Unbound protein was washed out with 3 column volumes (CV) PBS/ 2 mM EDTA/ 2 mM EGTA. Then, the envelope protein was eluted with PBS/ 2 mM EDTA/ 2 mM EGTA/ 1 M α-D-mannopyranoside (Merck Millipore, # 462711-100GM). Protein containing fractions were pooled, concentrated to a volume of 0.8 to 1 ml (Merck Millipore, Amicon Ultra 15 ml, 100K, # UFC910096) and Env was purified to trimer size homogeneity using a HiPrep 16/60 Sephacryl S-300 HR size exclusion chromatography (SEC) column (GE Healthcare, # GE17-1167-01) in PBS. Trimer containing fractions were pooled and concentrated (Merck Millipore, Amicon Ultra 4 ml, 30K, # UFC803024). Purity of Env trimers was analyzed by Blue Native PAGE (Serva, # 43204.01) and subsequent Coomassie staining. Protein aliquots were frozen in liquid nitrogen and stored at -80 °C.

Coomassie Brilliant Blue R-250	AppliChem (# A1092)
Coomassie staining solution	1.25 % (w/v) Coomassie Brilliant Blue R-250, 50 % (v/v) ethanol, 7 % (v/v) acetic acid

3.4.3. Chemical cross-linking of gp140 trimers

Cross-linking procedures used in this thesis were guided by published protocols.[78,79,50] For glutaraldehyde cross-linking (GLA, Sigma Aldrich, glutaraldehyde solution, 25 % in H_2O, # G5882), GLA and envelope trimers were diluted in PBS, mixed at ratios of 100, 1000, 2000, 3000, 4000, 5000 or 6000 mol GLA per mol envelope and incubated for 5 min at RT. To stop the reaction, 1 M Tris/HCl pH 7.4 was added to a final concentration of 75 mM, followed by 5 min incubation at RT. For EDC/NHS cross-linking ([(1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride/N-hydroxysuccinimide]), Sigma Aldrich, # 03449-5G and Thermo Fisher Scientific, # 24500), EDC and NHS were mixed at a molar ratio of 100 to 1 in MES (2-(*N*-morpholino)ethanesulfonic acid) buffered saline, pH 6.0. Subsequently, EDC/NHS and envelope were mixed at ratios of 35 000, 100 000, 160 000, 400 000 or 800 000 mol EDC per mol Env. Notably, the EDC/NHS mixture and envelope diluted in PBS were mixed using equal volumes. After incubation for 30 min at

RT, the reaction was stopped by adding a volume of 1 M glycine pH 7.4 equal to the volume of the total cross-linking reaction for 10 min at RT. Control gp140 envelopes for both cross-linking reactions received the same treatment without addition of cross-linking reagents. Success of cross-linking reactions was determined with reducing SDS-PAGE (5% separation gel, 5% stacking gel) as well as direct and capture ELISA immediately after cross-linking. If cross-linked envelopes were not instantly analyzed, buffer exchange to PBS was performed using Amicon Ultra 0.5 columns (Sigma Aldrich, # Z740174-96EA). To this end, the cross-linking reaction was diluted with approximately 450 µl PBS and centrifuged at 4000 g and 4 °C until the initial sample volume was reached. This procedure was conducted three times.

EDC stock solution	4 M N-(3-Dimethylaminopropyl-)-N'-ethylcarbodiimide hydrochloride (Sigma Aldrich, # 03449-5G) in MES buffered saline
NHS stock solution	5 M N-Hydroxysuccinimide (ThermoFisher Scientific, # 24500) in MES buffered saline
MES buffered saline	50 mM MES monohydrate, 150 mM NaCl), pH 6.0

3.4.4. SDS-PAGE and silver staining after cross-linking

Cross-linked envelope trimers were analyzed via SDS-PAGE with a 5 % stacking gel and 5 % separation gel. To this end, the protein samples were mixed with 5x Laemmli buffer, incubated for 5 min at 95 °C and shortly centrifuged (1 min, 5000 g). Electrophoresis was performed at 70 V for 20 min, followed by 90 V until the tracking dye reached the bottom of the gel.

Gels were subsequently silver-stained. Each step of this procedure was carried out using 25 ml of the respective solution at RT on an orbital shaker. First, fixing solution was applied for at least one hour (or overnight), followed by one wash step each with 50 % ethanol and 30 % ethanol for 10 min. Subsequently, the gel was incubated in sodium-thiosulfate (0.2 g/l) for 1 minute, followed by three wash steps for 20 seconds in H_2O. To stain the proteins, staining solution was added for 20 min, followed by two wash steps with H_2O for 20 sec. Next, the gel was incubated with developing solution until the protein bands were well visible or to a maximum of 10 min. Then, stop solution was added for 5 min. Until documentation, developed gels were stored in H_2O.

5x Laemmli buffer	312.5 mM Tris, 5 % SDS (20 %), 25 % β-mercaptoethanol, 25 % glycerol (87%), 2.5 mM EDTA (0.5 M, pH 8.0), pH 6.8 (HCl)
PageRuler[TM] Plus Prestained Protein Ladder	Thermo Fisher Scientific (# 26620)
Fixing Solution	50 % (v/v) methanol, 12 % (v/v) acetic acid, 50 µl pro 100 ml formaldehyde
Staining solution (100 ml)	0.1 % (w/w) $AgNO_3$, 75 µl 37 % formaldehyde
Developing solution (100 ml)	6 % (w/w) Na_2CO_3, 50 µl 37 % formaldehyde, 200 µl 0.2 g/l sodium thiosulfate
Stop solution	5 % (v/v) acetic acid

3.4.5. ELISA

ELISA plates (Nunc Maxisorp, # 442404) were coated with 200 µl of 1 µg/ml *Galanthus nivalis* lectin (Sigma, # L8275-5MG) in PBS overnight at 4 °C. All wash cycles were conducted with PBS/ 0.05 % Tween20 using a Tecan HydroFlex (HydroControl-Software Version 1.0.1.0). After coating, plates were washed three times and blocked for 2 h at RT with 200 µl 5 % (w/v) skimmed milk powder/ 5 % (v/v) heat inactivated FCS (Gibco)/ 0.1 % Tween20. After three further wash cycles, 50 µl of gp140 Env variants at a concentration of 1 µg/ml in PBS were added for 1 h at RT, followed by six wash cycles. Subsequently, 50 µl of single concentrations or a dilution series of human monoclonal envelope antibodies in PBS/ 1 % BSA (w/v) were added for 1 h at RT to reach equilibrium binding. 5F3, which recognizes a linear epitope in gp41[46] and should therefore recognize all variants alike, was used to monitor equal Env amounts for each variant on each plate. Unbound antibody was washed away with 10 wash cycles and bound antibody was labeled with 50 µl of horseradish peroxidase-coupled rabbit anti-human IgG (Agilent, Dako # P021402-2, 1:5000 in PBS/ 1 % BSA) for 30 min at RT, followed by 10 wash cycles and incubation with TMB substrate (TMB A: TMB B 20:1, composition see table below). The reaction was stopped with 25 µl 1 M H_2SO_4 and optical density at 450 nm (OD450) was analyzed in triplicates on a microplate reader.

As sCD4 caused unspecific signals with lectin coated plates, 70 µl of gp140 envelope at a concentration of 3 µg/ml in PBS were directly immobilized on the Maxisorp plates at 4 °C overnight for all sCD4 ELISAs. After three wash cycles, plates were blocked for 2 h at RT as described above. After three additional wash cycles, 50 µl of sCD4 (Progenics, Tarrytown; NY, USA, 1 µg/ml) in PBS + 1 % BSA were added and incubated for 1 h at

RT, followed by six wash cycles. Next, 50 µl of a 1:2500 dilution of α-human CD4 antibody (Biolegend, OKT4, # 317402) in PBS + 1 % BSA were added for 1 h at RT, followed by six wash cycles and subsequent incubation with 50 µl of rabbit α-mouse IgG-HRP (Dako, # P016102-2 , 1:1000 in PBS + 1 % BSA) for 30 min at RT. After six wash steps, incubation with TMB substrate and all following steps were conducted as described above. Capture ELISAs for analysis of cross-linking reactions were conducted as described for the lectin capture ELISA above, but with 2 µg/ml envelope instead of 1 µg/ml. For direct ELISAs after cross-linking, 2 µg/ml envelope dilution in PBS were directly coated onto the Maxisorp plates overnight at 4 °C or at least 2 h at RT. After three wash cycles, blocking and all subsequent steps were conducted as described above.

All curves were corrected by subtracting both the negative control signals obtained without envelope gp140 addition and without primary antibody and fitted using non-linear least squares regression (hyperbolar one site binding), GraphPad Prism 5.0 or 6.0). Area under the curve (AUC) values were also calculated in GraphPad Prism 6.0.

Wash program	Extraction of fluids
	Dispension of wash buffer (PBS/ 0.05 % Tween20, 300 µl/s)
	Wash cycles: 400 µl/s
	3 wash cycles: 1 ml wash buffer
	6 wash cycles: 2 ml wash buffer
	10 wash cycles: 3 ml wash buffer
	Extraction of fluids
TMB A	30 mM potassium citrate, pH 4.2
TMB B	10 mM tetramethylbenzidine (Roth 6350.2), 10 % (v/v) acetone, 90 % (v/v) ethanol, 80 mM H_2O_2

3.4.6. Antibodies

Human monoclonal antibodies HGN194 and HJ16 were kindly provided by Prof. Dr. Antonio Lanzavecchia and Dr. Davide Corti, Institute for Research in Biomedicine, Bellinzona, Switzerland. Antibodies PG9, 2G12, 4E10, 5F3 and 447-52D were purchased from Polymun Scientific, Klosterneuburg, Austria. VRC01, F105, PGT145, PGT151 and 10E8 were purified in house. To this end, 300 ml FreeStyle[TM] 293-F cells (Thermo Fisher Scientific, # R79007) were cultivated to a concentration of 1 Mio/ml. Directly before transfection, cells were centrifuged at 100 g for 5 min and resuspended in FreeStyle

expression medium (Thermo Fisher Scientific, # 12338018) without supplements. 6.5 ml DMEM medium without supplements were mixed with 300 µg plasmid DNA (150 µg light chain containing plasmid + 150 µg heavy chain containing plasmid). Additionally, another 6.5 ml DMEM without supplements were mixed with 1.5 ml PEI (Polyethylenimine, linear, Polysciences, # 23966-2, 1 mg/ml). After 5 min at RT, the two mixtures were combined and further incubated for 20 min at RT. Subsequently, the transfection mixture was given onto the cells. The cells were incubated at 37 °C, 8 % CO_2 and 90 rpm for six hours. Then, the medium was replaced with FreeStyle medium supplemented with 0.5 % Pen/Strep. 5 days post transfection, cells were harvested at 1000 g and 4 °C for 10 min. The supernatant was loaded onto a Protein A column (0.3 g Protein A sepharose (Sigma Aldrich, # P3391-1.5G) quelled with 5 ml buffer A (0.02 M NaH_2PO_4, 0.15 M NaCl, pH 8.0) for 30 min. Three column volumes of buffer A were used to wash out unbound protein, then the elution of IgG was carried out with 4 ml 100 mM glycine, pH 2.5. Eluted antibody was neutralized with stepwise addition of 1 M Tris/HCl, pH 8.0 until a pH of 6 to 8 was reached. For antibodies with pI values near 7, e.g. PGT145, neutralization was stopped at pH 6.0 to avoid antibody aggregation. Subsequently, antibodies were dialyzed against PBS (Thermo Fisher Scientific, Slide-A Lyzer dialysis cassettes 3.5K MWCO, # 66330). In the case of PGT145, the pH of PBS was adjusted from 7.4 to 6.0 to avoid antibody aggregation. Antibody concentration was determined using a NanoDrop 1000 (Thermo Scientific) and, if necessary, IgGs were concentrated using Amicon Ultra concentrators (Merck Millipore, 30 kD MWCO, # UFC903024). Antibody purity was analyzed by SDS-PAGE. All other antibodies were provided by the NIH Aids Reagents Program or IAVI (International AIDS Vaccine Initiative). sCD4 was provided by Progenics (Tarrytown, NY, USA) as indicated by the NIH AIDS reagents program.[100]

For FACS analysis, screening antibodies were labeled with Alexa Fluor 647 dye (Thermo Fisher Scientific, Protein labeling kits, # A-20173), while reference antibody 5F3 was labeled with R-Phycoerythrin (R-PE, Innova Biosciences, Lightning-Link® R-Phycoerythrin Conjugation Kit, # 703-0015). Soluble CD4 (Progenics, Tarrytown, NY, USA) was detected using a secondary APC-labeled anti-human CD4 antibody (anti-human CD4 APC, OKT4, eBiosciences, # 17-0048-42, 1:100).

3.4.7. Differential scanning calorimetry (DSC) of soluble Env trimers

Env trimers were dialyzed three trimes against PBS at 4 °C. Subsequently, dialyzed trimers as well as the final dialysis buffer were vented. Carefully avoiding the insertion of air into the sample, Env sample volumes were adjusted to 1 ml and a final concentration of 0.75 to 1.5 μM using degassed dialysis buffer. The protein sample was then applied to the cell of a Microcal VP Differential Scanning Calorimeter, while the vented dialysis buffer was applied to the reference cell. Thermal denaturation was performed at 1 °C/min from 20 to 95 °C and melting temperatures were calculated after normalization and baseline correction of the acquired data.

3.4.8. Nano differential scanning fluorimetry (nanoDSF) of soluble Env trimers

NanoDSF measurements were performed in cooperation with 2bind, Regensburg, Germany using a Prometheus NT.48 (NanoTemper). Briefly, 10 μl of gp140 trimers diluted to a concentration of 0.05 mg/ml in PBS were applied to Prometheus NT.48 series nanoDSF grade standard capillaries (NanoTemper) and measured with a temperature ingredient of 1 °C/min between 21 °C and 95 °C. A PBS negative control was used to exclude fluorescent signals caused by buffer components.

3.4.9. Surface plasmon resonance (SPR) of soluble Env trimers

SPR analysis for the comparison of labeled and unlabeled VRC01 and PG9 was carried out by the CAVD Central Service Facility at Duke University, Durham, USA. Briefly, the binding of 96ZM651 gp140 to Alexa Fluor 647-conjugated or unconjugated antibody of interest was measured on a Biacore 3000 (BIAcore/GE Healthcare). 96ZM651 gp140 was immobilized on a CM5 chip using standard amine coupling chemistry and serial dilutions of antibodies were injected. Alternatively, antibodies were captured with anti-human IgG

Fc antibody on a CM5 chip and a dilution series of 96ZM651 gp140 was used as soluble analyte. After each binding cycle, surfaces were regenerated with a 12 sec injection of 10 mM phosphoric acid (H_3PO_4). All measurements were carried out in triplicates. Data analyses were performed with BIAevaluation 4.1 software (BIAcore/GE Healthcare) and non-specific binding of antibodies to a negative control IgG surface was subtracted from each curve.

4. Results

4.1. Analysis of a mammalian cell display-based panning system for the high-throughput discovery of new AIDS vaccine candidates

Protection against SHIV infection after passive immunization of macaques with bnAbs[101,63,62] offered an important starting point for the development of an efficient HIV vaccine. Recent years have brought significant progress in envelope immunogen design (see 1.5.3). However, Env`s high variability,[18,26,27] extensive glycosylation,[29,102] conformational masking of target sites[30,31] and conformational instability[18,22] all represent effective immune escape strategies which hinder an efficient humoral immune response (see 1.3.3). Thus, the elicitation of broadly neutralizing antibodies by active immunization is still a major obstacle in HIV vaccine design. This emphasizes the need for innovative screening systems able to identify new Env vaccine candidates according to their antigenic profiles. Prior to this thesis, a mammalian cell display and flow cytometry-based panning procedure was developed by Dr. Tim-Henrik Bruun (Bruun and Grassmann *et al*., in revision), which enables the selection of Env variants with modulated antibody affinity. Here, this system was further characterized using an already established five-member chimeric Env model library[96] and human monoclonal antibodies (mAbs) HGN194[44] and 447-52D[103]. Analysis of enrichment rates was extended from qPCR and capillary sequencing of single clones to also include next generation sequencing (NGS) for the screening of larger envelope libraries.

4.1.1. Overview of the mammalian cell display-based panning system

The mammalian cell display and panning procedure comprises six different steps. It starts with the generation of stable cell lines with single integrations of one given Env variant per cell at a defined locus (linked cellular Env library, **Figure 11**). To this end, FlpIn^{TM} T-REx^{TM} 293 cells are transfected with an Env library cloned into the plasmid vector pQL13 (**Appendix Table 2** and **Figure 13A**), which supports the integration of a single copy Env gene into the predefined Flp Recombination Target (FRT) site of a FlpIn cell (**Figure 11**, step 1). Stable integration of pQL13 into the FRT site results in the acquisition of a hygromycin resistance (**Figure 13A**). Hence, after about 25 days of antibiotic selection, each of the resulting cells represents one Env variant, leading to the genotype-phenotype linked cellular Env library (**Figure 11**, step 2). Notably, Env expression is under regulation of the Tat operator/repressor system and thus inducible. After induction with doxycycline, Env expressing cells are stained with an antibody of interest (screening antibody, **Figure 11**, step 3) and variants can be selected by flow cytometry-based cell sorting according to their antibody binding profile (e.g. high affinity to a bnAb or low affinity to a non-neutralizing antibody) (**Figure 11**, step 4). To determine the genotypes of the selected Env variants, the genomic DNA of sorted cells and input controls is isolated and the contained Env genes are amplified by nested-PCR (**Figure 11**, step 5). After cloning into pQL13 and further propagation in *E. coli* (**Figure 11**, step 6), plasmid DNA of both the input controls and the output samples can be used to monitor the enrichment status of the different Env variants via qPCR, capillary sequencing of single clones or, in the case of larger libraries, next generation sequencing. Alternatively, genomic DNA of cells isolated prior to and after FACS-sorting can also be directly subjected to NGS, facilitating analyses of larger libraries. If necessary, selected cells can also be re-expanded for another round of panning to increase enrichment rates with each cycle (not shown in Figure 11). The different steps of this procedure are described in more detail in the following sections.
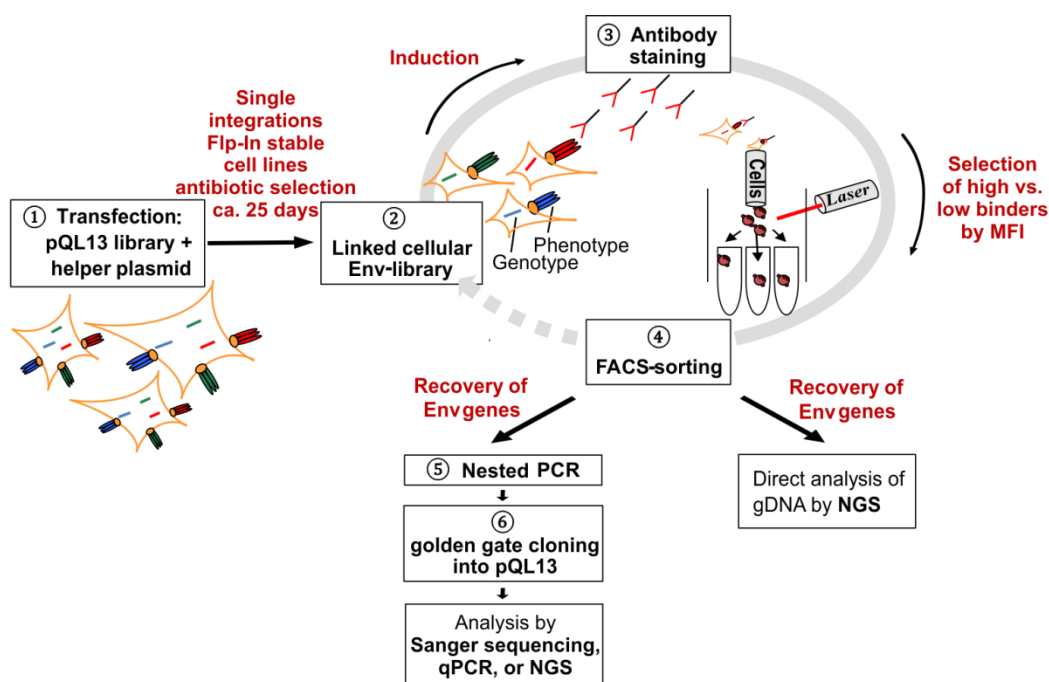
Figure 11: Schematic overview of the FACS panning procedure. Stable cell lines are generated by a targeted transfection of FlpIn™ T-REx™ 293 cells with pQL13 based Env constructs and the helper plasmid pOG44 carrying the integrase (1). Since FlpIn™ T-REx™ 293 cells possess precisely one FRT site, a linkage between geno- and phenotype is achieved. Successful integration into the FRT site leads to the acquisition of a hygromycin resistance, allowing antibiotic selection of the stable cell lines (2). After induction of Env expression, the cellular library is stained with the screening antibody (3) and subsequently, envelopes with the desired phenotype regarding antibody binding are selected via flow cytometry-based cell sorting (4). The genomic DNA of selected cells and input controls is then recovered and Env genes contained within are amplified by nested-PCR (5). Amplified envelope genes are cloned back into pQL13 and analyzed by qPCR, Sanger sequencing of single clones or next generation sequencing (6). Alternatively, gDNA recovered from the input sample and the sorted cells can directly be subjected to NGS to facilitate enrichment rate analysis for larger libraries. To enhance enrichment rates, a portion of the sorted cells can be re-expanded for another panning cycle (indicated by the broken line).

4.1.2. Generation of single-integration stable cell lines using a pQL13 based chimeric Env/V3 model library

The data referred to in paragraph 4.1.2 were generated by Dr. Tim-Henrik Bruun prior to this thesis, but are shown to better illustrate the panning procedure. Five chimeric gp145 Env variants were constructed based on a clade C 96ZM651 gp145 backbone. The V3 loop of this scaffold was seamlessly substituted with V3 regions of isolates MN, RF, CDC42, HXB2 or SF33 (**Figure 12**), which were known to exhibit distinguishable affinities towards the V3-targeting mAb 447-52D.[96,103] Using the "golden gate" cloning procedure,[96]

the model library was then cloned into pQL13. The toxic CcdB cassette in the MCS of the pQL13 target vector thereby allows for an efficient selection of envelope positive clones to facilitate the library assembly in *E. coli*.
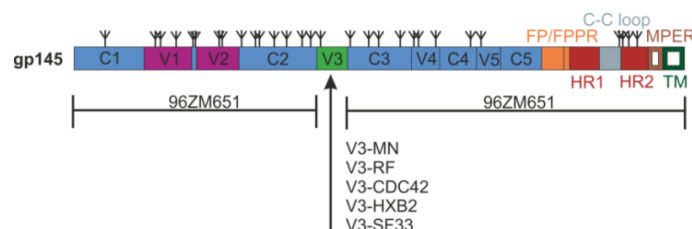
Figure 12: Schematic representation of the chimeric Env/V3 model library used in the panning approach. V3 of the common clade C 96ZM651 gp145 scaffold was seamlessly substituted with the V3 regions of isolates MN, RF, CDC42, HXB2 or SF33, resulting in five chimeric variants. See Figure 4 for all other abbreviations in this schematic.

The pQL13-Env/V3 model library was used to generate stable cell lines by targeted transfection of FlpIn[TM] T-REx[TM] 293 cells with the helper plasmid pOG44 carrying the integrase and (i) an equimolar mixture of all five Env/V3 chimeras (cell library) or (ii) each Env/V3 chimera individually. While the cell library was applied in all panning experiments, the individual cell lines were used to characterize the antigenic profile of each chimera. All stable cell lines displayed several unique features:

(i) Single integration of one Env variant per cell. Integration of pQL13 into the predefined FRT site at a specific locus (**Figure 13A**) ensured the presence of only one Env variant per cell, efficiently linking genotype and phenotype within the library. The integration of a single pQL13 copy per cell for all scenarios tested was shown with a TaqMan[®] copy number assay probing for eGFP in relation to the human telomerase reverse transcriptase (TERT) genes (**Figure 13B**).

(ii) Antibiotic selection. Cells with a successful integration event could be selected via a hygromycin resistance, which is acquired and constitutively expressed only after recombination into the FRT site, when the hygromycin gene lacking an ATG comes into contact with a start codon (**Figure 13A**, ca. 25 days of selection necessary).

(iii) Translational coupling of eGFP and Env expression. Translational coupling of eGFP and Env by a "self-cleaving" TaV 2A peptide[96,104] (**Figure 13A**) enabled normalization for Env expression according to eGFP fluorescence (MFI). The insect virus-derived (*Thosea asigna* virus) 2A peptide impairs peptide bond formation by a ribosomal

skip mechanism between the 2A glycine and the 2B proline of the consensus sequence 2A,Asp-Val/Ile-Glu-X-Asn-Pro-Gly; 2B,Pro. This should result in an equal expression level for eGFP and Env, while only one residue is N-terminally attached to the envelope protein. To investigate the correlated expression of eGFP and Env, doxycycline-induced cells were stained with mAb 5F3,[105] which recognizes a linear epitope in the extracellular domain of gp41 present in all five Env/V3 variants. A correlation between eGFP and Env expression was observed for each chimera (**Figure 13C**), verifying eGFP as an indirect marker to normalize for Env expression. Notably, expression rates of the variants differed (with Env/V3-MN displaying the weakest signals), highlighting the importance of expression normalization.

(iv) Inducible Env expression. Finally, the stable integration of the Tet repressor in the FlpIn[TM] T-REx[TM] cells in combination with the doxycycline-inducible CMV promotor of pQL13 results in the regulated expression (T-REx) of the integrated Env gene (**Figure 13A**), allowing cultivation of the stable cell lines without negative effects potentially caused by Env cytotoxicity. Notably, the T-REx system only utilizes regulatory elements from the native Tet operon,[106] avoiding the potentially toxic effects of viral transactivation domains utilized for other doxycycline-regulated systems.[107]
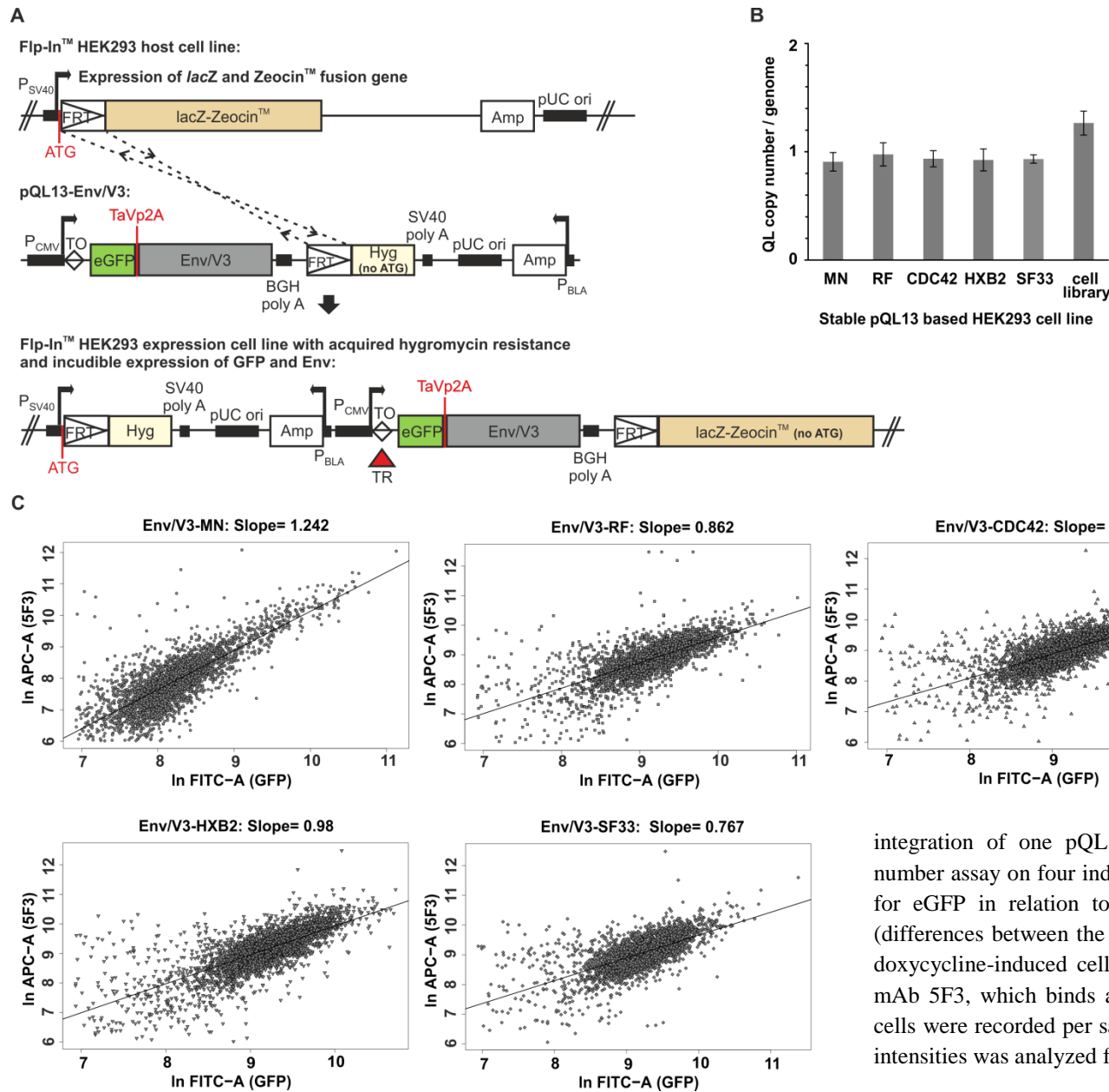
Figure 13: Characteristics of the pQL13-based stable cell lines. Stable cell lines were generated using separate transfection of each pQL13-Env/V3 chimera or an equimolar mixture of all five Env/V3 variants (cell library). **A:** Schematic overview of the pQL13 vector system before and after stable integration into FlpIn[TM] T-REx[TM] 293 cells. In pQL13, Env expression is genetically linked to eGFP expression by a TaV 2A peptide (TaVp2A). The eGFP/Env expression cassette is integrated into a distinct Flp Recombination Target (FRT) site in the FlpIn[TM] T-REx[TM] 293 cells, bringing the hygromycin gene in contact with a start codon and hence resulting in the acquisition of a hygromycin (Hyg) resistance for selection of stable cell lines. Expression of eGFP and Env is under inducible regulation of the Tet operator/repressor (TO/TR, T-REx[TM]). Note that pQL13 is a circular plasmid which is here represented by its horizontal vector map and that the symbols for the FRT site (52 bp) and the TO (9 bp) are not in scale. **B:** Single integration of one pQL13 per genome could be verified using a TaqMan copy number assay on four individual samples of gDNA for each cell line and probing for eGFP in relation to the human telomerase reverse transcriptase (TERT) (differences between the six stable cell lines statistically n.s., $p>0.05$). **C:** 3×10^5 doxycycline-induced cells of each individual stable cell line were stained with mAb 5F3, which binds a constant region shared by all variants (n=2). 2.5×10^4 cells were recorded per sample and the correlation between APC and GFP signal intensities was analyzed for each stable cell line (3000 events shown).

4.1.3. The Env/V3 model library exhibits distinct binding profiles to the monoclonal antibodies 447-52D and HGN194

Using the individual stable cell line of each Env/V3 chimera, the binding profiles of V3-targeting mAbs 447-52D and HGN194 to each variant were determined by FACS equilibrium titration (**Figure 14B** and **D**). Env/V3-MN proofed to have the highest affinity towards 447-52D and HGN194, followed by chimeras –HXB2 and –CDC42. Whereas HGN194 did bind neither Env/V3-RF, nor -SF33, 447-52D could still bind Env/V3-RF to some extent. For these non- or only weakly binding variants, K_D values could not be calculated. For all other variants, K_D values are summarized in **Table 2**. **Figure 14A** and **C** show the distinct binding of 447-52D and HGN194 to the five chimeras at the respective antibody concentration used in the panning procedure. In conclusion, the demonstrated distributions of affinities were considered feasible for the purpose of evaluating the stable cell line-based panning procedure.
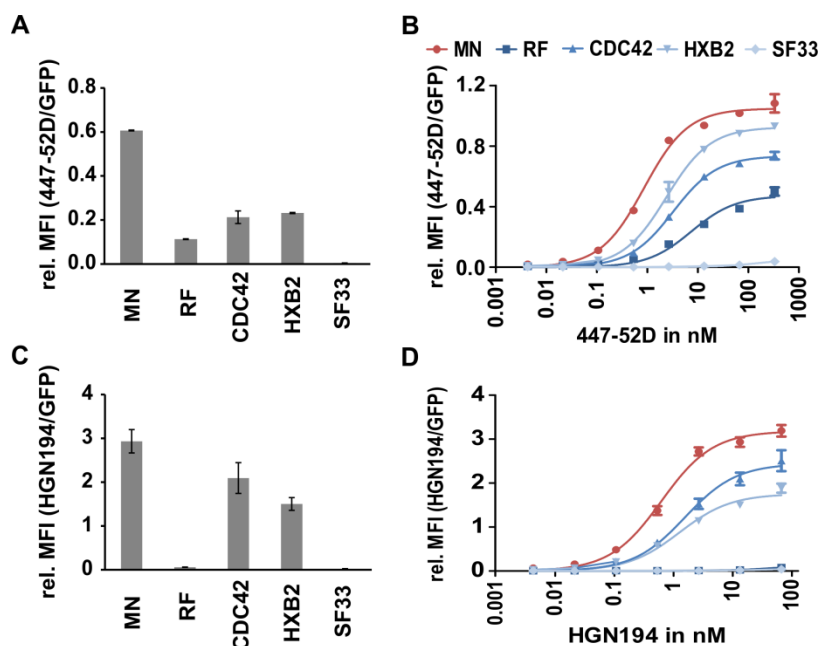
Figure 14: Affinity analysis of the individual Env/V3 gp145 chimeras to 447-52D and HGN194. Induced stable cell lines were equilibrium stained with mAbs 447-52D or HGN194 (n=3). Relative MFI values of antibody binding in relation to GFP were calculated to normalize for Env expression levels. **A** and **C**: Cells were stained with 2.7 nM 447-52D (A) or 6.7 nM HGN194 (C) as used in all FACS-sorting experiments. **B** and **D**: Serial dilutions of 447-52D (B) or HGN194 (D) were applied to obtain a concentration-dependent binding profile and calculate K_D values (Table 2) (n=3). Chimera Env/V3-MN depicted the highest affinity to 447-52D and HGN194, followed by -CDC42 and -HXB2. 447-52D still showed weak binding to chimera Env/V3-RF but did not recognize -SF33. HGN194 did bind neither -RF nor -SF33.

4.1.4. Single-round FACS panning of the Env/V3 cell library leads to selective isolation of high or low affinity variants

The cell library generated with an equimolar mixture of all five Env/V3 chimeras was subjected to flow cytometry-based cell sorting using approximately 3×10^7 cells per sort. Triangular gates were chosen to sort cells with the highest (P5) or lowest (P7) 447-52D or HGN194 signal in relation to GFP, thus selecting cells with the highest or lowest antibody affinity relative to Env expression (**Figure 15**).
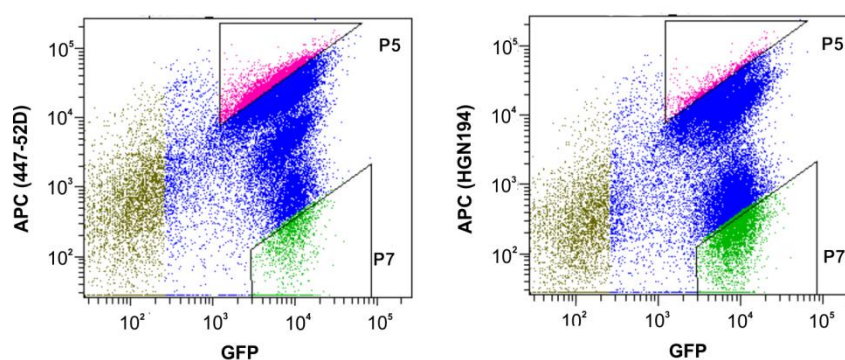
Figure 15: Gating strategy of the panning procedure. A typical sorting experiment is shown (50,000 events recorded). Living, single cells were gated according to common hierarchical gating strategies (gates P1-4, not shown). Cells split up into uninduced cells (low GFP/ low APC) and induced cells with a higher GFP signal. The latter can be further divided into cells with high APC signals (high affinity antibody binding) and cells with low APC signals (low affinity antibody binding). Induced cells were gated for highest (P5) or lowest (P7) APC signals (447-52D or HGN194) in relation to GFP (expression control), resulting in triangular gates. P6 represents a gate for the sorting of medium affinity binders between P5 and P7 (not shown). P5 was set to contain 600 – 800 of the 50,000 recorded cells, while P7 was chosen to contain approximately half of the induced population with low APC signal intensity.

The genomic DNA from an input sample and the sorted cells (~30,000 cells/sorting gate) was recovered and Env genes were amplified by nested-PCR and cloned back into pQL13. The distribution of Env/V3 variants in the input control and after one cycle of panning was monitored by qPCR, capillary sequencing of single clones and NGS using the pQL13 plasmid DNA recovered from *E. coli*. In addition, NGS was performed directly on the gDNA to provide a facilitated analysis approach for larger libraries (**Figure 16**).

In the case of 447-52D, a statistically significant enrichment of high affinity Env/V3 variant -MN out of an equal distribution of all variants was demonstrated for each analysis

approach (**Figure 16A**). Thereby, enrichment rates of up to 56-fold could be achieved after only a single round of panning (calculation **Table 1**). Gate P7 selectively enriched low affinity variant Env/V3-SF33 (**Figure 16A**), with enrichment factors of up to 237-fold (**Table 1**). Hence, this panning technology can similarly be used to isolate Env variants with particularly low binding to unwanted antibodies, e.g. non-neutralizing mAbs targeting the CD4-induced Env conformation. In summary, the MFIs measured at a distinct 447-52D antibody concentration (**Figure 14A**) as well as binding affinities of the individual Env/V3 chimeras (**Figure 14B**, **Table 2**) correlated with the copy number (qPCR) or number of sequences (Sanger sequencing, NGS) quantified after the single-round FACS panning process. These findings could also be confirmed with the independent HIV-1 Env V3-specific monoclonal antibody HGN194 (**Figure 16B**). Here, high affinity variant -MN could be enriched up to 55-fold fold after one round of panning (**Table 1**). Similarly, as chimeras Env/V3-RF and -SF33 both did not show binding to HGN194 (**Figure 14C** and **D**, **Table 2**), low signal gate P7 selectively enriched those two variants with enrichment factors of up to 4-fold for chimera -RF and up to 5-fold for chimera -SF33 (**Figure 16B**, **Table 1**).

Notably, results from the already established Sanger sequencing and qPCR analyses were conform with the newly established NGS approach using the Illumina Miseq platform on both the re-cloned plasmid DNA as well as directly on the genomic DNA. While NGS will enable the panning of very large libraries, the direct analysis of recovered gDNA will additionally avoid the difficulties originating from re-cloning a high number of different variants. Variations in enrichment factors between the different analyses (**Table 1**) may be caused by diverging inherent properties of the different methods and the sensitivity of the applied formula, which considers both the enriched chimeras as well as the depletion of all other variants. As NGS supplies a very high sequence coverage for the here applied five-member model library, these findings might be the most accurate.
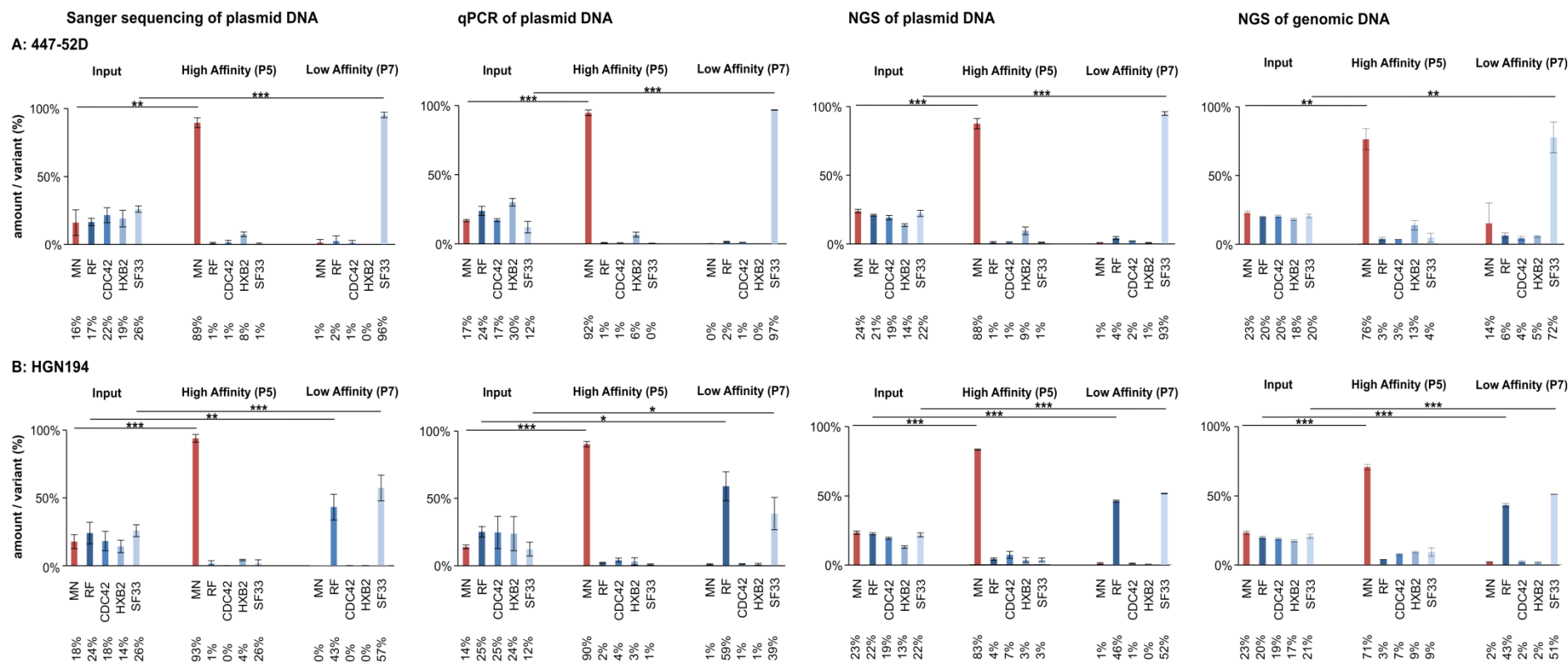
Figure 16: Single-round FACS panning using the Env/V3 cell library. 24h after doxycycline induction, the stable cell line containing the complete Env/V3 model library (cell library) was used in the FACS panning procedure with **A:** 2.7 nM 447-52D or **B:** 6.7 nM HGN194 (n=3). Envelope genes selected after one panning cycle as well as from the input control were amplified from the genomic DNA, cloned back into pQL13 and plasmid DNA recovered from *E. coli* was analyzed by qPCR, Sanger sequencing of single clones and next generation sequencing using the Illumina MiSeq platform. In addition, NGS was performed directly on the genomic DNA. Relative amounts per variant in the input mixture and after the 1st round of panning are shown in %. Statistically significant enrichment of the high affinity variant Env/V3-MN was achieved for both 447-52D and HGN194 (P5). Similarly, gate P7 led to selective enrichment of the low affinity variants (Env/V3-SF33 for 447-52D and both Env/V3-RF and -SF33 for HGN194). Statistics were calculated using an unpaired t-test in GraphPad Prism 5. *=p<0.05, **=p<0.01 and *** p < 0.001.

Table 1: Enrichment rates of Env/V3 chimeras selected from the FACS-sorting procedure after one round of panning.

		Fold enrichment[a] in high affinity gate P5	Fold enrichment[a] in low affinity gate P7	
		MN	RF	SF33
447-52D	qPCR (pDNA)	56	No enrichment	237
	Sanger (pDNA)	43	No enrichment	68
	NGS (pDNA)	23	No enrichment	41
	NGS (gDNA)	11	No enrichment	10
HGN194	qPCR (pDNA)	55	4	5
	Sanger (pDNA)	14	2	4
	NGS (pDNA)	16	3	4
	NGS (gDNA)	9	3	4

a) calculations were made according to the formula from [108]

$$\left(\frac{MN\ \#1}{MN\ \#0} \Big/ \frac{Non\text{-}MN\ \#1}{Non\text{-}MN\ \#0}, \frac{RF\ \#1}{RF\ \#0} \Big/ \frac{Non\text{-}RF\ \#1}{Non\text{-}RF\ \#0} \text{ or } \frac{SF33\ \#1}{SF33\ \#0} \Big/ \frac{Non\text{-}SF33\ \#1}{Non\text{-}SF33\ \#0} \right)$$

4.1.5. Soluble gp140 Env/V3 chimeras display a similar binding pattern to 447-52D and HGN194 as membrane-bound variants.

The five soluble Env/V3 gp140 chimeras were purified to trimer homogeneity and affinities towards 447-52D and HGN194 were analyzed by ELISA equilibrium titration (**Figure 17**). This resulted in very similar binding patterns compared to the gp145 Env-displaying cell lines (**Figure 14**) and dissociation constants (K_D) were found to be in the same order (**Table 2**), with Env/V3-MN displaying the highest affinity, followed by -CDC42 and -HXB2. Also in agreement with the membrane-bound variants, 447-52D still showed weak binding to Env/V3-RF, while HGN194 did neither bind Env/V3-RF nor -SF33. For those low affinity binders, K_D values were not calculable.
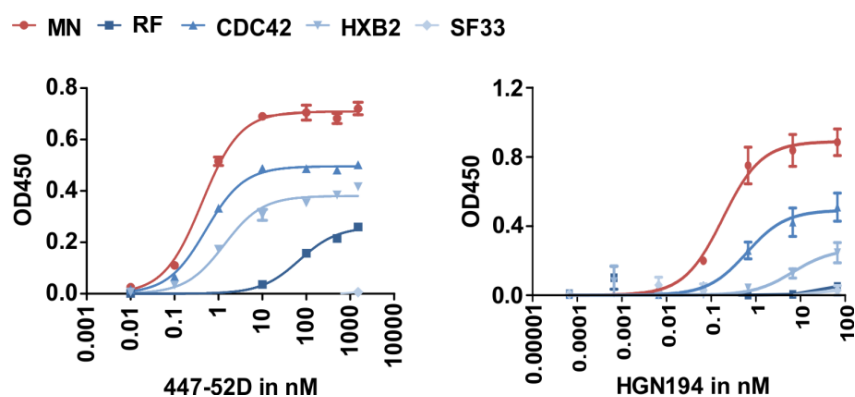
Figure 17: ELISA titration of soluble gp140 Env/V3 chimeras. Trimeric gp140 envelopes were purified and equilibrium titrated with 447-52D or HGN194 via ELISA (n=3, one biological replicate exemplary shown). The soluble trimers depicted the same 447-52D and HGN194 binding patterns as membrane-bound gp145 variants (Figure 14).

Table 2: Dissociation constants (K_D) of membrane-bound gp145 chimeras obtained by FACS equilibrium titration and soluble gp140 chimeras obtained by ELISA equilibrium titration. N=3. R^2 values of the respective fits (hyperbolar one site binding, GraphPad Prism 5.0) are included in brackets. Due to low antibody binding, K_D values of Env/V3-SF33 to both 447-52D and HGN194 and of Env/V3-RF to HGN194 were not calculable. K_D values of V3 peptides to 447-52D, which served as a selection criteria for the different V3 loops, are taken from [103]. Membrane-bound and soluble Env/V3 chimeras of the model library displayed the same affinity ranking.

Env/V3 chimera	447-52D			HGN194	
	K_D [nM] of V3 peptides	K_D [nM] of membrane-bound gp145 (flow cytometry titration, n=3)	K_D [nM] of soluble gp140 (ELISA titration, n=3)	K_D [nM] of membrane-bound gp145 (flow cytometry titration, n=3)	K_D [nM] of soluble gp140 (ELISA titration, n=3)
MN	0.6	0.9 ± 0.1 (R^2=0.990)	0.4 ± 0.02 (R^2=0.995)	0.6 ± 0.04 (R^2=0.981)	0.2 ± 0.05 (R^2=0.926)
RF	0.9	7 ± 1 (R^2=0.972)	74 ± 7 (R^2=0.992)	NA	NA
CDC42	5	3 ± 0.2 (R^2=0.996)	1 ± 0.4 (R^2=0.983)	1 ± 0.2 (R^2=0.973)	0.6 ± 0.2 (R^2=0.797)
HXB2	24	2 ± 0.2 (R^2=0.991)	0.5 ± 0.002 (R^2=0.983)	2 ± 0.09 (R^2=0.951)	7 ± 5 (R^2=0.537)
SF33	NA	NA	NA	NA	NA

4.1.6. Summary

In sum, the FACS panning procedure was able to enrich high and low affinity binders from a cell surface-displayed envelope library, as expemplified with the five-member model library and mAbs 447-52D and HGN194. Analysis of the membrane-bound and soluble Env/V3 variants demonstrated a clear correlation between (i) the MFIs measured at a distinct 447-52D or HGN194 antibody concentration for the individual Env/V3 chimeras (**Figure 14A** and **C**), (ii) their binding affinities to 447-52D or HGN194 (**Table 2**), (iii) the copy number (qPCR) or number of sequences (capillary sequencing or NGS) of the respective Env/V3 chimeras after sorting (**Figure 16**) and (iv) binding affinities of 447-52D or HGN194 to the soluble Env/V3 trimers (**Figure 17**, **Table 2**). Notably, the enrichment analysis of the sorting procedure could successfully be adapted to enable the efficient screening of large libraries.

4.2. Development of HIV-1 envelope immunogens with improved antigenicity and stability using a cell-based epitope mapping platform

Understanding Env − bnAb interactions in detail is valuable for the rational design of next generation Env immunogens aiming at a more potent and cross-neutralizing antibody response *in vivo*. However, these reverse vaccinology approaches are highly dependent on structural and functional data regarding epitope characterization. The required information is often obtained from Env/antibody co-crystallization[109,110], binding studies with envelope mutants[111,54,112] or peptide scans[44,113]. Here, a cell surface antibody mapping (CSAM) platform was utilized to characterize envelope - bnAb interactions and identify mutations for the rational design of novel Env immunogens. To this end, an Env alanine library was applied in a mammalian cell display-based high-throughput flow cytometry assay, enabling native folding and mammalian glycosylation as well as display of trimeric Env in its natural membrane context.

4.2.1. General setup of the epitope mapping procedure

The envelope library as well as the general procedure of the CSAM platform was developed by Dr. Alexander Kliche prior to this thesis. An alanine mutagenesis library based on the C clade isolate 96ZM651 was generated to address the globally most prevalent subtype C. To ensure efficient Env expression and surface presentation, the sequence was adapted to human codon usage[114] and the cytoplasmic tail was truncated[64,65]. Alanine mutants were generated from the signal sequence C-terminus to the N-terminus of the transmembrane domain (**Figure 18**, Library Design). Cysteine residues were excluded from mutagenesis to preserve Env structure, resulting in a total of 605 variants. In addition, the furin protease cleavage site was mutated from REKR to REKS to prevent processing of the precursor protein into the non-covalently linked gp120-gp41 complex prone to gp120 shedding.[115,116]

The gp120 mutants of this library were applied in high-throughput 96-well transfection of HEK293T cells (one variant per well) and Env expressing cells were analyzed by flow

cytometry to assess each amino acid mutation for its impact on antibody binding. To this end, the antibodies used in this procedure were covalently labeled with a fluorescent dye (Alexa Fluor 647 or R-Phycoerythrin (R-PE)), resulting in a direct correlation between antibody binding and the fluorescence signal intensity and allowing the simultaneous use of more than one mAb from the same species. 48 h after transfection, each well was simultaneously stained with the screening antibody and reference antibody 5F3. This antibody recognizes a linear epitope in the C-terminal heptad region of gp41[46] and should therefore bind all gp120 alanine variants alike. Hence it was used to normalize for varying Env surface expression levels and transfection efficiencies (**Figure 18**, Separate Transfection and Antibody Staining). MFI ratios for the screening and the reference antibody were calculated and all gp120 mutants were analyzed for significantly altered binding levels in comparison to the WT reference (**Figure 18**, High-Throughput FACS Analysis).

As a first proof of concept, Dr. Alexander Kliche analyzed the binding of HGN194, a neutralizing mAb with a linear epitope previously mapped to the crown of the V3 loop.[44] A profound and statistically significant reduction of HGN194 binding was only observed for variants I307A, I309A, G310A and Q313A (data not shown, numbering according to HXB2). These alanine mutants only showed between 5 and 26 % binding capacity for HGN194 compared to the 96ZM651 wildtype and were previously published by Corti *et al.* to be critical for HGN194 binding in pep-scan analysis[44].

In the course of this thesis, the CD4 binding site-targeting bnAb VRC01 was screened, as well as quaternary structure-dependent, apex-targeting bnAb PG9. In addition, a soluble CD4 (sCD4) alanine scan previously performed by Dr. Alexander Kliche was newly analyzed and significant mutations were further utilized.
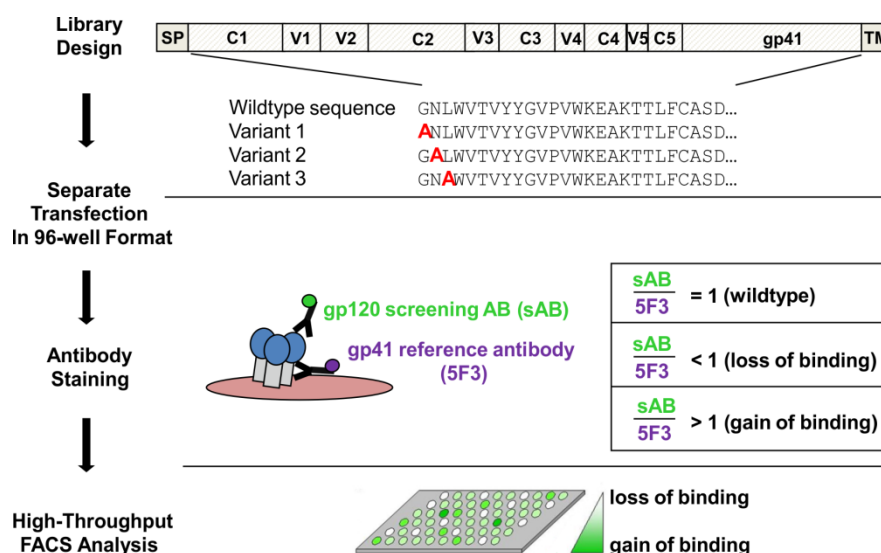
Figure 18: Schematic representation of the cell surface antibody mapping platform. A 96ZM651 gp145 (clade C) alanine scanning library was generated. The gp120 alanine mutants were used for separate 96-well transfection of HEK293T cells (one variant per well). 48 h post transfection, Env expressing cells were stained with the screening antibody and a gp41 reference antibody (5F3), which recognizes a linear epitope present in all variants, allowing normalization for varying transfection efficiencies and Env surface expression levels. Stained cells were analyzed by flow cytometry and MFI ratios of the screening antibody and the reference antibody were analyzed for significantly altered antibody binding in comparison to the wildtype. SP: signal peptide, C1-C5: constant regions of gp120, V1-V5: variable loops of gp120, TM: transmembrane domain.

4.2.2. Evaluation of the alanine scanning procedure for sCD4, VRC01 and PG9

For all screening reagents of this thesis (i.e. sCD4, CD4 binding site-targeting bnAb VRC01 and apex-targeting bnAb PG9), the following parameters of the CSAM procedure were evaluated: (i) use of cleavage-incompetent gp145 REKS for library generation, (ii) direct labeling of the applied antibodies and (iii) simultaneous staining with the screening and reference antibody.

(i) Validation of 96ZM651 cleavage-incompetent (REKS) gp145 for the mapping of sCD4, VRC01 and PG9. The binding of sCD4, VRC01 and PG9 to the 96ZM651 cleavage-competent REKR wildtype as well as to the cleavage-defective REKS mutant used for library generation was compared in the flow cytometry-based assay format described above. The cleavage-defective 96ZM651 gp145 showed no difference in binding

to sCD4, VRC01 and PG9 compared to cleavage-competent Env, even when cleavage was supported by co-transfection of furin (**Figure 19**), validating the use of the cleavage-incompetent REKS library for these reagents.
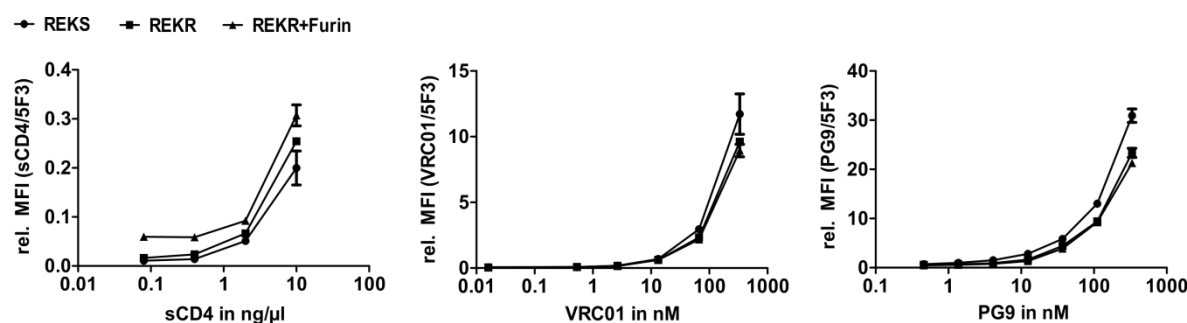
Figure 19: Impact of cleavage site mutation on the binding of sCD4, VRC01 and PG9. HEK293T cells were transfected with (i) the cleavage-defective (REKS) gp145 which served as a basis for the alanine scanning library, (ii) 96ZM651 cleavage-competent (REKR) 96ZM651 gp145 and (iii) 96ZM651 cleavage-competent (REKR) gp145 and furin to support cleavage. 48 h post transfection, Env expressing cells were stained with sCD4, VRC01 or PG9. To normalize for Env expression levels, 5F3 co-staining was applied. Neither sCD4, nor VRC01 and PG9 showed significantly altered binding to the cleavage-defective variant in comparison to 96ZM651 gp145 wildtype (REKR), even when cleavage was supported by co-transfection of furin. n=2.

(ii) Direct labeling of antibodies for the CSAM procedure. While sCD4 could be detected by an anti-human sCD4 secondary antibody in this assay setup, the envelope antibodies of interest for epitope mapping as well as the reference antibody 5F3 are of human origin. This prevents the use of an anti-human IgG secondary antibody for detection. Hence, antibodies applied in the CSAM procedure were directly labeled with a fluorescent dye (Alexa Fluor 647 for screening antibodies or R-PE for reference antibody 5F3). To analyze the effect of this untargeted labeling reaction on the binding capacity of VRC01 and PG9, the binding of unmodified vs. labeled antibodies to 96ZM651 gp140 WT was compared by surface plasmon resonance (SPR). In addition, antibody functionality was analyzed against a small number of different isolates. Labeling reduced VRC01 binding capacity approximately 2.5 to 3-fold in two different SPR settings (**Figure 20A and C**). In accordance with these results, VRC01-Alexa647 also showed about threefold decreased activity in the TZM-bl neutralization assay (**Table 3**). PG9-Alexa647 also exhibited slightly reduced binding compared to the unmodified mAb (**Figure 20B and D),**

although in this case antibody functionality was not influenced (**Table 3**). The untargeted labeling reaction used in the CSAM procedure covalently attaches four to six dye molecules per antibody molecule based on a covalent linkage of the succinimidyl ester group of the fluorescent dye to the primary amines of the antibody. Amino groups in the antibody paratope reacting with the fluorescent dye may therefore lead to blockage of the paratope in a certain percentage of antibody molecules. In the case of VRC01, a lysine residue at position 52 in the heavy chain forms a hydrogen bond to Ala281 of gp120.[42] For PG9, crystal structures with V1/V2 peptides indicate an interaction of different heavy and light chain lysines with Env.[117] However, blockage of the antibody paratope would only lead to a diminished effective concentration of applied antibody, therefore Alexa647 labeled mAbs were still deemed suitable for epitope mapping.
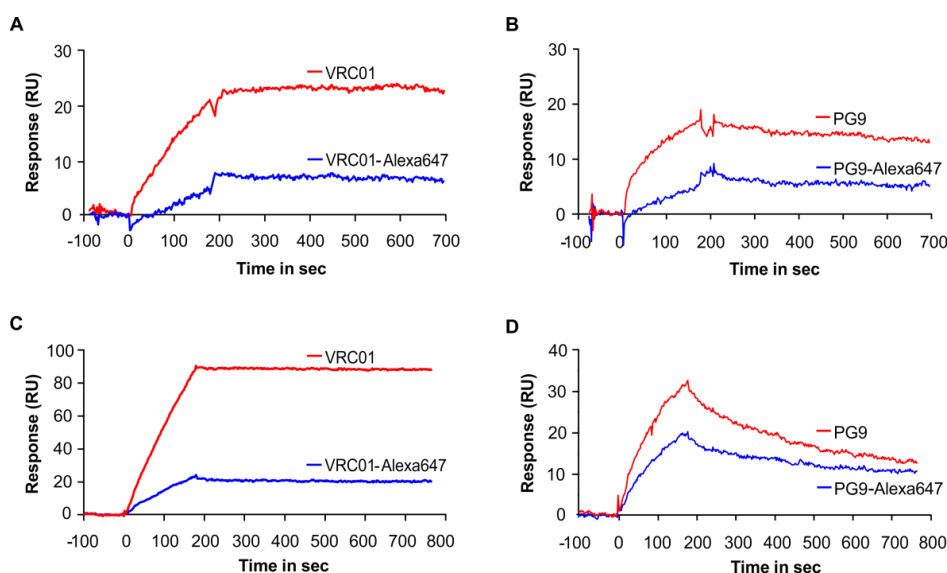
Figure 20: Impact of direct fluorescent labeling on antibodies VRC01 and PG9. mAbs VRC01 and PG9 were labeled with fluorescent dye Alexa Fluor 647 and binding of unmodified and labeled antibodies to 96ZM651 gp140 was analyzed by SPR. **A** and **B:** Antibodies were captured on a CM5 chip using anti-human IgG Fc antibody and 96ZM651 gp140 was applied as soluble analyte. **C** and **D:** 96ZM651 gp140 was directly immobilized on a CM5 chip using amine coupling and mAbs VRC01 and PG9 were used as soluble analytes. RU: resonance units. Labeling reduced gp140 binding of both VRC01 and PG9.

Table 3: Impact of direct fluorescent labeling on VRC01 and PG9 functionality. VRC01-Alexa647 showed weaker neutralization capacity than the unmodified mAb in a TZM-bl assay against a panel of different pseudoviruses, while PG9 functionality was not affected.

Virus	Clade	IC50 Titers in TZM-bl cells (µg/ml)			
		VRC01	VRC01-Alexa647	PG9	PG9-Alex647
Q23.17	A	0.09	0.33	0.00	0.00
Q842.d12	A	0.03	0.09	0.03	0.02
REJO4541.67	B	0.09	0.30	0.01	0.01
TRJO4551.58	B	0.08	0.31	0.61	0.40
WITO4160.33	B	0.10	0.38	0.01	0.01
SC422661.8	B	0.14	0.55	0.60	0.96
AC10.0.20	B	1.49	4.67	0.11	0.12
ZM249M.PL1	C	0.10	0.31	0.06	0.06
ZM53M.PB12	C	0.62	3.42	0.06	0.05
ZM197M.PB7	C	0.42	2.22	0.54	0.63
Murine Leukemia Virus	Neg. Control	>50	>50	>50	>50

(iii) Simultaneous staining of Env expressing cells with the screening reagent and reference antibody 5F3. MFI values of the screening compounds sCD4, VRC01 and PG9 were comparable after single staining and simultaneous staining with 5F3 (**Figure 21A**), indicating that no sterical hindrances arise from the double staining procedure. Furthermore, a correlation between signal intensities of 5F3 and each screening compound was observed for gp145 wildtype transfected cells (**Figure 21B**, **Table 4**). Several statistical parameters were collected to evaluate assay quality for each screening reagent (**Table 4**). As standard deviations must always be understood in relation to the mean of the data and mean values varied considerably for the different compounds, the coefficient of variation, also known as relative standard deviation, was calculated as a standardized measure of data point dispersion. Coefficients of variation were found to be 26 % for reference antibody 5F3, while all other screening compounds reached values under 20 %. Confidence intervals (p = 0.05) were low, ranging from 2.9 % (VRC01) to 5.5 % (5F3). This indicates that e.g. for 5F3, 95 % of the measured values would not differ more than 5.5 % from the mean. In addition, the Z´-value was calculated, which describes an excellent assay with 0.5< Z´<1, a doable assay with 0<Z´<0.5 and Z´=0 as a yes/no type of assay.[118,119,120] While the Z´-value was not largely influenced by 5F3 normalization in the case of VRC01, Z´ could be improved from 0.35 to 0.57 for sCD4. The most profound effect of 5F3 normalization was observed for PG9, where an improvement from 0.27 to 0.58 was reached. In comparison to the other screening compounds, PG9 depicts weak

affinity to 96ZM651 envelope and only showed a signal to background ratio of 3 (**Table 4**). Hence, this alanine scan was more error-prone than the others and might therefore have benefitted the most from normalization. In summary, the setup of the screening procedure was deemed suitable for the mapping of sCD4, VRC01 and PG9 binding sites.
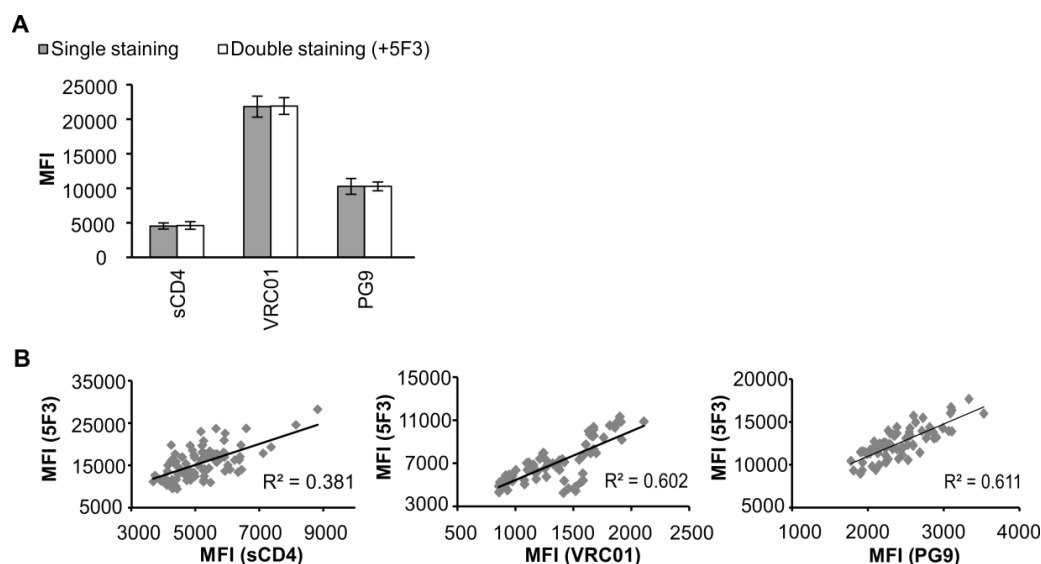
Figure 21: 5F3 as a reference antibody to normalize for Env surface presentation. **A:** To analyze the influence of simultaneous staining with a screening compound and 5F3, 293T cells were transiently transfected (n=10) with 96ZM651 gp145 REKS. 48 h post transfection, cells were either stained with the indicated interaction partner alone (grey bars) or in combination with antibody 5F3 (white bars) and analyzed by flow cytometry. No sterical hindrances could be detected for the simultaneous staining approach. **B:** A 96-well plate of 293T cells was transfected with 96ZM651 gp145 REKS and mean MFIs of each well were analyzed for correlation between the signals of reference antibody 5F3 and the respective screening compound (n=80).

Table 4: Statistical parameters of the FACS-based alanine scanning procedure. 293T cells were transfected with 96ZM651 gp145 REKS in a 96-well format (n=80, negative control: n=6) and simultaneously stained with one of the screening compounds (sCD4, VRC01 or PG9) and 5F3 to be analyzed by flow cytometry. Several statistical parameters were calculated to describe assay performance for each interaction partner. SD: standared deviation. Based on low confidence intervals, coefficients of variation and Z´-values > 0.5, the CSAM procedure was deemed suitable for the purpose of epitope mapping.

	Non-normalized values of the screening compounds						Values normalized by reference antibody 5F3					
	Mean MFI	SD	Confidence interval, p=0.05	Signal to background ratio	Coefficient of variation (SD/Mean)	Z´	Mean MFI	SD	Confidence interval, p=0.05	Coefficient of variation (SD/Mean)	Z´	R^2 (screening AB / 5F3)
sCD4	5168	925	4977 – 5369 or ± 3.7%	40	18%	0.35	0.340	0.065	0.327 – 0.354 or ± 3.9%	19%	0.57	0.381
VRC01	2061	294	2000 – 2121 or ± 2.9%	15	14%	0.61	0.196	0.043	0.186 – 0.205 or ± 2.0%	22%	0.53	0.602
PG9	2411	378	2333 – 2489 or ± 3.4%	3	16%	0.27	0.193	0.020	0.189 – 0.197 or ± 4.0%	10%	0.58	0.611
5F3	12484	3305	11802 – 13167 or ± 5.5%	29	26%	0.19						

4.2.3. Analysis of envelope interactions with soluble CD4

4.2.3.1. Alanine scan of soluble CD4

Dr. Alexander Kliche analyzed the well-characterized interaction of Env and sCD4[20,111,121] to validate the CSAM method for complex discontinuous epitopes. All shown analyses of this screening were newly conducted by me. The sCD4 scan identified 159 alanine mutations which resulted in significantly reduced binding compared to the parental envelope. 48 out of those were located within the inner domain[20] of gp120 and are thus buried in the native trimer, while 70 mutations were located in the outer domain[20]. Of the latter, 29 were in the variable loops V1/V2 and one in V3. In addition, six significant mutations were located in the bridging sheet (**Figure 22**).
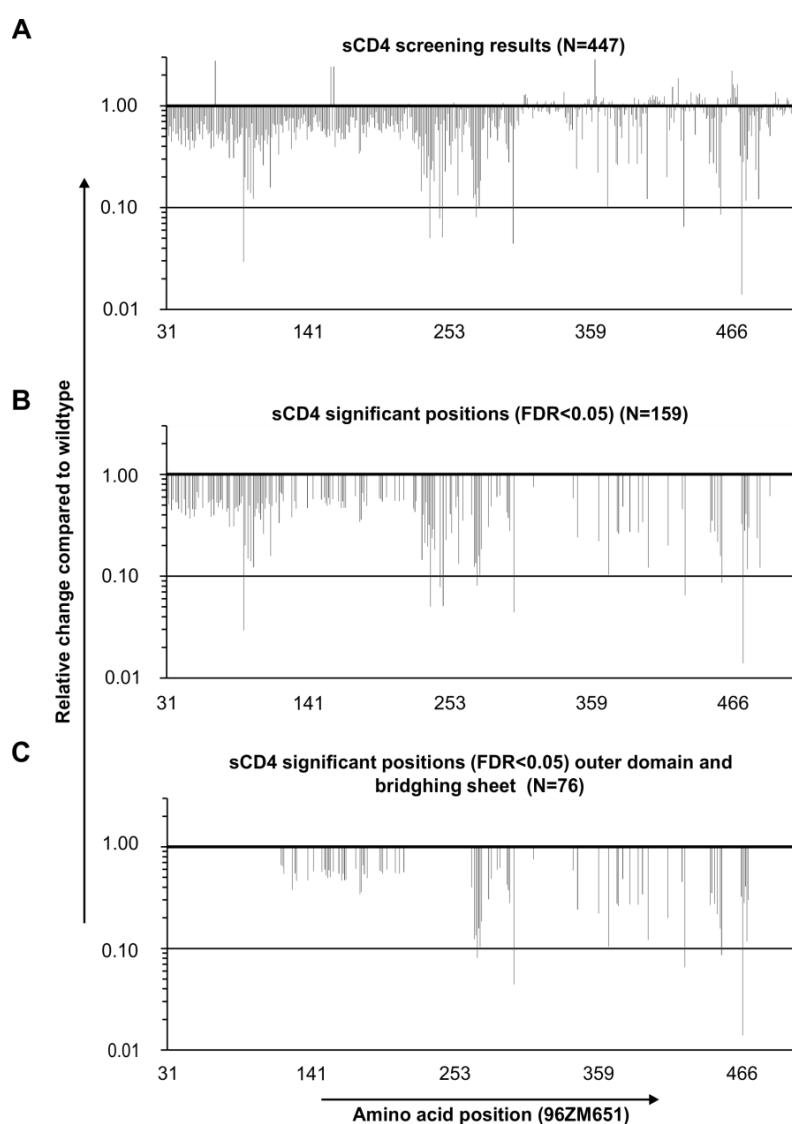
Figure 22: Alanine scan of soluble CD4. 5F3 normalized fold changes of sCD4 binding relative to 96ZM651 WT (WT≙1) are shown for every gp120 mutant of the alanine scanning library (n=6). The diagrams show **A:** all positions screened with the CSAM method (N=447), **B:** all positions which showed statistically significant fold changes (FDR<0.5, N=159) and **C:** all positions which reached significance and are located in the outer domain of Env and in the bridging sheet (N=76) and are thus accessible from the surface. Statistical significance was determined in R using a two-sided one sample t-test against 1, while controlling the false discovery rate (FDR) to be smaller than 5 % to adjust for multiple testing.

All statistically significant mutations are illustrated on the crystal structure of BG505 SOSIP.664 (PDB ID: 4ZMJ)[122] and compared to the CD4 contact residues defined by Kwong *et al.* (**Figure 23**).[20,42] In summary, Dr. Kliche could identify several CD4 contact residues by significantly reduced binding of sCD4 to the generated alanine mutants (**Figure 23A**). However, not every contact site mutation could decrease sCD4 recognition. Especially in β20/β21 of the bridging sheet, only mutation of N425 and the highly conserved W427 had a severe impact on CD4 binding (**Figure 23A**).[123] Several reasons might account for this: In contrast to crystallography, which detects the whole interaction surface of two binding partners, this mapping technology only detects contacts that contribute to the interaction energy and are hence really necessary for binding.[124] On the other hand, the alanine scan only removes side chain interactions between Env and the ligand. If an alanine exchange does not lead to a significant alteration of backbone architecture, main-chain-only contacts would not be detected.[125,111] This, for example, could explain the negative influence of mutations N425A and W427A in the bridging sheet, where main chain and side chain contacts account for the interaction.[42] In contrast, neighbouring residues like M426, Q428 and G431 were not influenced, as only main chain interactions play a role at these positions (**Figure 23A**).[42] In summary, the general region of the epitope could be made apparent (**Figure 23C**).

In addition to residues with direct contact to sCD4, significantly reduced sCD4 binding could also be observed for other alanine substitutions, for example within the α1 helix (**Figure 23B**, black arrow). The proper packing of the α1 helix of the inner domain with the outer domain and the β20/β21 strands is critical for the formation of the four-stranded bridging sheet.[126] In addition, most of the additional residues which significantly influenced sCD4 binding, including those in α1, could also be observed for the structure-dependent antibodies VRC01 and PG9 (**Appendix Table 3**). Notably, these residues are located in the constant regions of Env and are, with very few exceptions, highly conserved. Hence, it is likely that they are important for correct Env folding.

Residue	FC	Contact
P124A	0.71	●
L125A	0.72	●
C126A	NA	●
V127A	0.57	●
C196A	NA	☼
T198A	0.71	☼
N279A	0.74	●
N280A	0.92	●
V281A	1.08	●
K282A	0.68	●
T283A	0.42*	●
S365A	0.95	●
G366A	0.64	○
G367A	0.27*	○
D368A	0.26*	●
L369A	0.91	☼
E370A	0.63	☼
H374A	0.91	☼
N425A	0.45*	●
M426A	0.91	●
W427A	0.07*	●
Q428A	0.88	○
E429A	1.14	●
V430A	NA	●
G431A	0.88	○
R432A	1.36	☼
T455A	1.20	☼
R456A	0.16*	○
D457A	0.09*	●
G458A	0.69	○
G459A	1.03	○
S460A	0.96	●
T461A	0.80	☼
R469A	0.28*	☼
G472A	0.12*	○
G473A	0.30*	○
D474A	0.57	●
M475A	0.81	☼
R476A	0.50	☼
D477A	0.88	☼

Figure 23: Structural analysis of the FACS-based sCD4 alanine scan. A: 5F3 normalized alanine scan fold changes (FC) relative to 96ZM651 WT are indicated for the published contact residues of gp120 (HXB2 numbering) with sCD4. Statistically significant (FDR<0.05) results are asterisked. Contact residues are taken from [42] with open circles (○) denoting gp120 main-chain-only contacts, open circles with rays (☼) denoting gp120 side-chain-only contacts, and filled circles (●) denoting both main chain and side chain contacts. **B:** Ribbon diagram of BG505 SOSIP.664 (PDB ID: 4ZMJ). Cyan: inner domain, blue: outer domain, green: bridging sheet, gray: gp41, black arrow: α1 helix. **C:** Surface representation of BG505 SOSIP.664 (PDB ID: 4ZMJ). For **B** and **C**, statistically significant amino acids (FDR<0.05) from the alanine scan are shown in red, CD4 contact residues as determined by crystallography[20,42] in yellow and the overlap between the two methods is depicted in orange.

4.2.3.2. Further analysis of N425 and W427 mutants

Binding of Env to its primary receptor CD4 initiates the profound conformational change necessary for host cell entry. Thereby, various epitopes for non-neutralizing antibodies usually hidden in the inner part of functional, closed trimers are revealed.[22] Hence, Env immunogens that do not bind endogenous CD4 would likely be less prone to induce non-neutralizing antibodies. Furthermore, decreased CD4 binding would improve bioavailability, as Env antigens could not be trapped on the surface of host $CD4^+$ immune cells.[79] To be included into next generation envelope immunogens, mutations which significantly decrease CD4 binding would additionally have to preserve bnAb recognition. Hence, after comparison with all available bnAb alanine scanning data in the group (data not shown), four of the significant sCD4 loss of binding (LOB) mutants were tested against

a panel of bnAbs targeting different epitopes. Two of those, variants N425A and W427A, showed profoundly reduced binding of sCD4 in the flow cytometry-based assay, while bnAb binding was preserved (**Table 5**). Notably, both mutations also led to severely reduced binding of non-neutralizing antibody 17b, which targets the CD4-induced (CD4i) conformation of Env (also seen in the 17b alanine scan, **Appendix Table 4**). Therefore, N425A and W427A represent valuable candidates for the rational design of new Env immunogens.

Table 5: Impact of N425A and W427A on different epitopes on 95ZM651 gp145. 5F3 normalized fold changes for the binding of 96ZM651 gp145 N425A and W427A to different mAbs and sCD4 as determined by flow cytometry are given (n=3, WT≙1). Both mutants showed profoundly reduced binding to sCD4 and 17b, while bnAb binding to different epitopes was preserved. MPER: membrane-proximal external region. CD4BS: CD4 binding site, CD4i: CD4-induced envelope conformation. Data were generated during the master thesis of Matthias Glögl under my experimental supervision.

| | | broadly neutralizing | | | | | | | non or weakly neutralizing | | |
| | Mutant | MPER | | Glycan | V3 | | CD4BS | Apex | CD4i | CD4BS | |
		4E10	10E8	2G12	HGN194	PGT121	VRC01	PG9	17b	b6	sCD4
membrane-bound gp145	N425A	1.33 ± 0.01	1.09 ± 0.09	1.70 ± 0.30	1.15 ± 0.05	0.85 ± 0.19	0.75 ± 0.08	1.01 ± 0.12	0.29 ± 0.05	0.75 ± 0.16	0.33 ± 0.04
	W427A	1.11 ± 0.08	1.10 ± 0.12	1.34 ± 0.51	1.08 ± 0.18	1.10 ± 0.16	0.81 ± 0.06	0.83 ± 0.22	0.14 ± 0.04	1.43 ± 0.26	0.07 ± 0.04

| 0 | 1 | 5 |

In addition, the sCD4 loss of binding effects could be transferred to gp145 envelopes from clades A, B and C. In agreement with the data obtained for 96ZM651, W427A thereby showed the more profound fold change and was hence also introduced into two SOSIP constructs, which exhibited the same effect (**Table 6**). This highlights that results from the 96ZM651 alanine scanning procedure can also be valuable for the rational design of Env immunogens from other clades or modern approaches like the SOSIP design.

Clade	Isolate	sCD4 FC N425A	sCD4 FC W427A
A	BG505	0.43 ± 0.09	0.08 ± 0.04
A	BG505 SOSIP	NA	0.03 ± 0.01
B	HXB2	1.05 ± 0.05	0.05 ± 0.01
B	JRFL	0.53 ± 0.17	0.04 ± 0.01
C	16055	0.33 ± 0.07	0.07 ± 0.02
C	16055 SOSIP	NA	0.11 ± 0.03
C	96ZM651	0.33 ± 0.02	0.07 ± 0.02

| 0 | 1 | 2 |

Table 6: Transfer of N425A and W427A to other isolates. gp145 mutants of clades A, B and C were analyzed by flow cytometry. Fold changes of 5F3 normalized sCD4 MFIs in relation to the respective WT are given (n=3). In agreement with the data obtained for 96ZM651, W427A reduced sCD4 binding more profoundly for all isolates tested and was therefore also introduced into BG505 SOSIP and 16055 SOSIP, yielding similar results. Data were generated during the master thesis of Christina Schmalzl under my experimental supervision.

4.2.4. Analysis of envelope interactions with CD4 binding site-targeting bnAb VRC01

4.2.4.1. Alanine scan of VRC01

The screening of all gp120 alanine variants with bnAb VRC01 resulted in 81 statistically significant changes in binding capacity. 66 of those were located in the outer domain of Env, whereof two residues were located in V3. 14 positions were located in the inner domain (**Figure 24**).
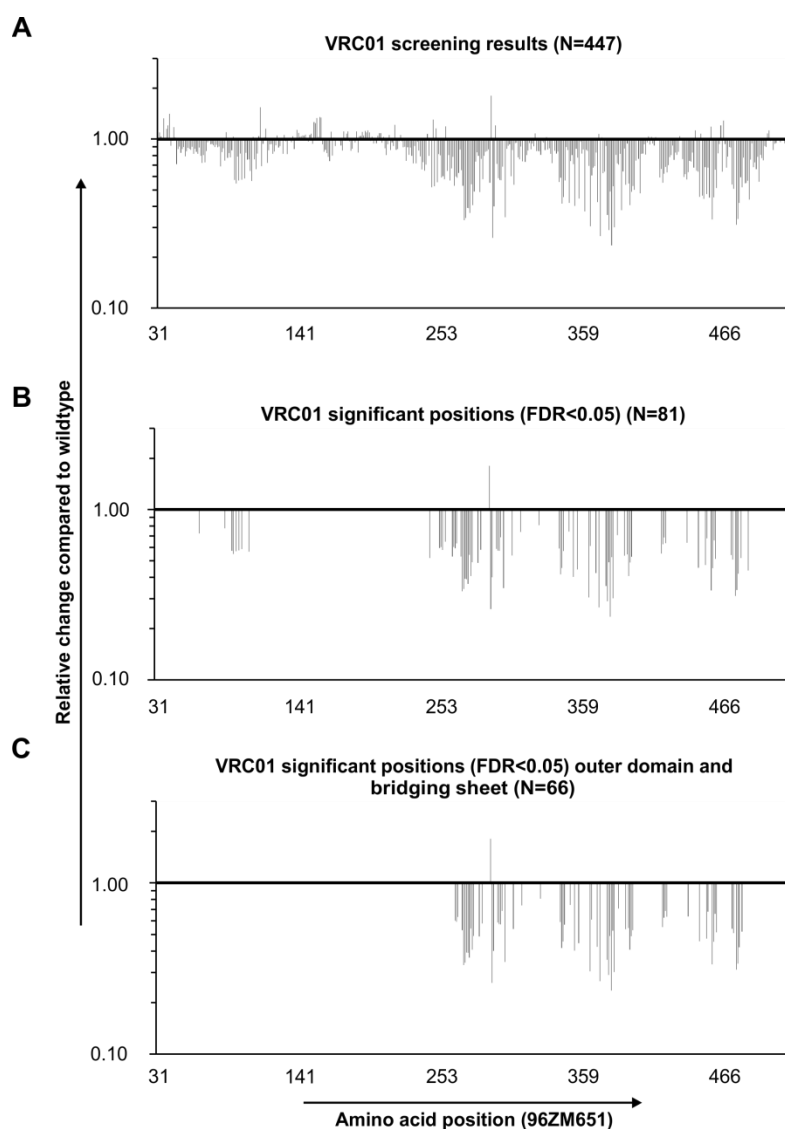
Figure 24: Alanine scan of VRC01. 5F3 normalized fold changes of VRC01 binding relative to 96ZM651 WT (WT≙1) are shown for every gp120 mutant of the alanine scanning library (n=6). The diagrams show **A:** all positions screened with the CSAM method (N=477), **B:** all positions which showed statistically significant fold changes (FDR<0.5, N=81) and **C:** all positions which reached significance and are located in the outer domain of Env and the bridging sheet (N=66) and are thus accessible from the surface. Statistical significance was determined in R using a two-sided one sample t-test against 1, while adjusting the false discovery rate (FDR) to be smaller than 5 % to correct for multiple testing.

Co-crystallization of VRC01 and Env reported 36 contact residues for this interaction.[42] The 96ZM651 alanine scan was able to identify 17 of those 36 contacts by a significant impact on VRC01 binding. Seven of the 19 amino acid positions that could not be confirmed are proposed to be main-chain-only contacts[42] (**Figure 25A**) and would therefore not be expected to alter VRC01 binding in this assay, as the backbone structure of Env in the respective area would likely not be influenced.[111,125] 80 of the 81 significant alanine mutations led to a severe loss of VRC01 binding, while one position, mutant T278A (HXB2 numbering), resulted in a 1.8 fold increase in VRC01 recognition (**Figures 24** and **25A**). This position is part of a crucial N-glycosylation site (N276 X277 T278) at the CD4 binding site that many VRC01-class bnAbs avoid by a CDR-L1 deletion. In addition, this glycan results in a steric clash for the germline version of VRC01 and has to be removed in order to allow germline binding.[91] This suggests that removal of this glycan leads to better accessibility of the CD4 binding site, thus facilitating VRC01 recognition. However, also N276A would remove this glycan, but with a fold change of 0.87 (data not shown), the alanine scan could not detect a significant influence ($p = 0.99$) of this mutation on VRC01 signal intensities. The beneficial effect of glycan removal by N276A could supposedly be abolished by a negative structural effect of this mutation on the CD4 binding site, as this mutant also showed significant reduction of sCD4 binding (fold change: 0.6, $p = 0.03$, **Appendix Table 4**), although N276 is not a contact residue for sCD4 (**Figure 23A**).[42]

A

Residue	FC	Contact
K97A	0.82	☼
T123A	0.92	○
T128A	1.01	○
L129A	0.82	○
N276A	0.87	☼
T278A	1.81*	●
N279A	0.26*	●
N280A	0.40*	●
V281A	1.21	●
K282A	0.71	☼
T283A	0.59*	☼
S365A	0.91	●
G366A	0.36*	○
G367A	0.29*	○
D368A	0.49*	●
I371A	0.30*	☼
W427A	0.81	○
Q428A	0.84	○
E429A	1.04	○
V430A	NA	☼
T455A	1.19	☼
R456A	0.33*	●
D457A	0.45*	●
G458A	0.66*	○
G459A	0.51*	○
N460A	0.83	●
S461A	0.94	●
N463A	0.79	☼
S465A	0.89	●
E466A	0.54*	●
I467A	0.51*	●
R469A	0.31*	☼
G472A	0.42*	○
G473A	0.60	○
D474A	0.52*	●
R476A	0.55	☼

0 — 1 — 2

B

C

D

Figure 25: Structural analysis of the FACS-based alanine scan of VRC01. A: Binding data of the 96ZM651 alanine scan for the published contact residues of VRC01 with gp120. The numbering of amino acids is based on HXB2. 5F3 normalized fold changes (FC) of VRC01 signals compared to wildtype binding levels are given and statistically significant results (FDR<0.05) are asterisked. Contact residues are taken from [42] with open circles (○) denoting gp120 main-chain-only contacts, open circles with rays (☼) denoting gp120 side-chain-only contacts, and filled circles (●) denoting both main chain and side chain contacts. **B:** Ribbon diagram of BG505 SOSIP.664 (PDB ID: 4ZMJ). Cyan: inner domain, blue: outer domain, light green: bridging sheet. Statistically significant amino acids (FDR<0.05) showing a loss of VRC01 binding in the alanine scan are depicted in red, the statistically significant gain of binding mutation in dark green (black arrow), VRC01 contact residues as determined by crystallography[42] in yellow and the overlap between crystal structure contact residues and the alanine scan is shown in orange. **C** and **D:** Surface representation of BG505 SOSIP.664 (PDB ID: 4ZMJ). The statistically significant gain of binding mutation is depicted in dark green. Statistically significant amino acids found to reduce VRC01 binding in the alanine scan are shown in red, VRC01 contact residues (C) or CD4 contact residues (D) identified by crystallography[42,20] in yellow and the overlap of amino acids between the alanine scan and crystal structure data is shown in orange.

4.2.4.2. Further characterization of T278 mutants

Position T278 was permutated to evaluate whether other amino acid substitutions at this site can lead to further improvement of VRC01 binding. Cysteine and proline likely disturbed Env structure and therefore resulted in reduced recognition of VRC01. In accordance with the hypothesis that glycan removal at N276 allows better access to the CD4 binding site,[91] most other amino acid substitutions resulted in increased VRC01 signals, with T278H depicting the most pronounced fold change (**Figure 26A** and **C**). This superior effect of the histidine mutant might be explained by the close proximity of position 278 to Y28 and Y91 of the VRC01 light chain, as a positive charge could strengthen the interaction with those VRC01 residues (**Figure 26B**). Also, as its conformational epitope renders VRC01 structure-dependent, the additional benefit of T278H compared to T278A might be caused by a positive influence of the histidine mutant on Env tertiary structure. K_D values of VRC01 binding to membrane-bound 96ZM651 could not be reliably determined, as binding curves did not reach saturation (**Figure 26C**).

Notably, the effect of enhanced VRC01 binding for T278A and H in membrane-bound gp145 was also transferable to soluble 96ZM651 gp140. While 96ZM651 gp140 WT showed a K_D value of 38.7 nM to VRC01 in ELISA measurements, T278A and T278H depicted K_D values of 16.1 and 11.4 nM, respectively (**Table 8**). Interestingly, T278H not only positively influenced VRC01 binding, but also resulted in increased affinity of apex-targeting bnAb PG9, accounting for approximately threefold increased PG9 signals for membrane-bound T278H compared to 96ZM651 WT and a K_D improvement from 62.0 to 30.3 nM for soluble gp140 (**Tables 7** and **13**). As PG9 is a quaternary structure-preferring antibody[127] and T278A did not affect PG9 affinity (**Table 13**), this further supports the hypothesis that T278H improves Env folding.

Other bnAb epitopes on Env could be preserved for both membrane-bound and soluble 96ZM651 variants (**Table 7**).
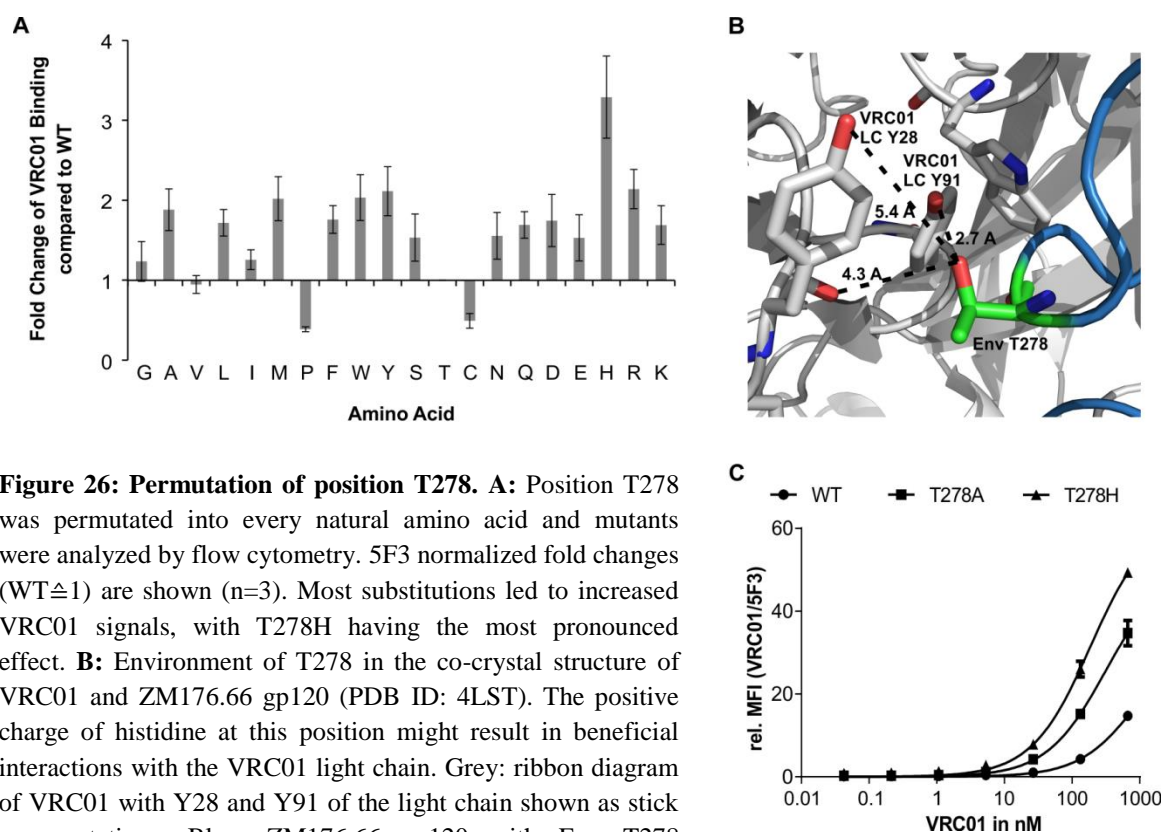
Figure 26: Permutation of position T278. A: Position T278 was permutated into every natural amino acid and mutants were analyzed by flow cytometry. 5F3 normalized fold changes (WT≙1) are shown (n=3). Most substitutions led to increased VRC01 signals, with T278H having the most pronounced effect. **B:** Environment of T278 in the co-crystal structure of VRC01 and ZM176.66 gp120 (PDB ID: 4LST). The positive charge of histidine at this position might result in beneficial interactions with the VRC01 light chain. Grey: ribbon diagram of VRC01 with Y28 and Y91 of the light chain shown as stick representations. Blue: ZM176.66 gp120 with Env T278 illustrated as stick representation and coloured in green. **C:** 5F3 normalized FACS titration of 96ZM651 gp145 WT, T278A and T278H with VRC01 (n=2). K_D values could not reliably be determined, as saturation could not be reached.

Table 7: Impact of T278A and T278H on different epitopes of 96ZM651 membrane-bound gp145 and soluble gp140. Fold changes are indicated for gp145 as determined by flow cytometry (5F3 normalized) and gp140 trimer fractions as determined by ELISA, with utilized antibody concentrations determined from Appendix Figure 3 (n=3, WT≙1). Both gp145 and gp140 T278H depicted improved signals for VRC01 and PG9, while other bnAb epitopes were preserved. MPER: membrane-proximal external region, CD4BS: CD4 binding site, CD4i: CD4-induced conformation of Env. Data for T278A gp145 were generated during the master thesis of Matthias Glögl under my experimental supervision.

		broadly neutralizing							non or weakly neutralizing		
	Mutant	MPER		Glycan	V3		CD4BS	Apex	CD4i	CD4BS	
		4E10	10E8	2G12	HGN194	PGT121	VRC01	PG9	17b	b6	sCD4
membrane-bound gp145	T278A	0.96 ± 0.17	1.11 ± 0.13	1.32 ± 0.44	1.06 ± 0.12	1.10 ± 0.17	1.73 ± 0.51	1.20 ± 0.06	0.80 ± 0.27	0.93 ± 0.21	0.62 ± 0.27
	T278H	1.17 ± 0.15	1.07 ± 0.01	0.97 ± 0.06	1.10 ± 0.01	1.22 ± 0.03	5.32 ± 0.10	3.37 ± 0.26	1.35 ± 0.06	1.09 ± 0.01	1.43 ± 0.13
soluble gp140	T278A	0.87 ± 0.06	0.86 ± 0.10	0.75 ± 0.03	0.89 ± 0.06	0.67 ± 0.00	1.29 ± 0.02	1.13 ± 0.21	0.81 ± 0.08	0.89 ± 0.03	0.80 ± 0.05
	T278H	1.04 ± 0.00	1.02 ± 0.01	0.87 ± 0.33	1.01 ± 0.02	1.14 ± 0.07	1.72 ± 0.60	1.65 ± 0.13	0.96 ± 0.08	1.04 ± 0.01	0.96 ± 0.02

0 1 5

To test whether the benefits observed for 96ZM651 are transferable to other isolates, T278A and H were introduced into BG505 (clade A), HXB2 (clade B), JRFL (clade B) and 16055 (clade C), as well as into the trimer-stabilized constructs BG505 SOSIP and 16055 SOSIP. Flow cytometry analysis of the respective gp145 variants confirmed that both mutations led to increased recognition of VRC01 for all tested isolates, with T278H tending to exhibit the superior effect (**Table 8**). Therefore, all histidine mutants were purified as soluble gp140 trimers and K_D values to VRC01 were determined by ELISA titration, resulting in approximately two- to threefold higher affinity for all T278H variants compared to the respective WTs (**Table 8**). Notably, magnitudes of K_D values thereby varied between the isolates, with 96ZM651 depicting the lowest and BG505 the highest inherent affinity to VRC01.

Table 8: Binding of VRC01 to membrane-bound gp145 and soluble gp140 T278A and T278H mutants of different isolates. Env gp145 T278A and T278H mutants of isolates from clades A, B and C were analyzed for VRC01 binding by flow cytometry (n=3). 5F3 normalized fold changes (FC) are indicated for each variant relative to the respective WT. Both mutations led to increased VRC01 binding, with T278H having a slightly superior effect. Hence, T278H mutants were further analyzed as soluble gp140 trimers by ELISA (n=2), where K_D values were found to be reduced about threefold for all variants compared to gp140 WTs. Corresponding ELISA titration curves are shown in Appendix Figure 2. Data were generated during the master thesis of Christina Schmalzl under my experimental supervision.

Clade	Isolate	Variant	cell surface gp145	soluble gp140	
			FC of VRC01 binding to membrane-bound gp145	K_D of VRC01 binding to soluble gp140 in nM	FC of VRC01 binding to soluble gp140
A	BG505	WT	1	4.3 ± 0.4	1.0
		T278A	1.11 ± 0.03	NA	NA
		T278H	1.50 ± 0.15	1.8 ± 0.2	2.4
A	BG505 SOSIP	WT	1	12.3 ± 2.2	1.0
		T278A	1.69 ± 0.06	NA	NA
		T278H	2.86 ± 0.41	3.6 ± 0.7	3.4
B	HXB2	WT	1	11 ± 0.8	1.0
		T278A	2.19 ± 0.31	NA	NA
		T278H	2.42 ± 42	3.7 ± 0.4	3.0
B	JRFL	WT	1	6.8 ± 0.5	1.0
		T278A	2.14 ± 0.15	NA	NA
		T278H	2.20 ± 0.18	2.7 ± 0.3	2.5
C	16055	WT	1	12.2 ± 0.1	1.0
		T278A	1.81 ± 0.11	NA	NA
		T278H	1.76 ± 0.07	4 ± 0.2	3.1
C	16055 SOSIP	WT	1	13.5 ± 1.0	1.0
		T278A	1.81 ± 0.31	NA	NA
		T278H	2.02 ± 0.31	6 ± 0.6	2.3
C	96ZM651	WT	1	38.7 ± 2.7	1.0
		T278A	1.88 ± 0.11	16.1 ± 0.6	2.4
		T278H	3.29 ± 0.21	11.4 ± 0.6	3.4

| 0 | 1 | highest FC | highest K_D | lowest K_D | 0 | 1 | highest FC |

4.2.5. Analysis of envelope interactions with quaternary structure-dependent bnAb PG9

4.2.5.1. Alanine scan of PG9

PG9 is a quaternary structure-dependent antibody asymmetrically targeting V1/V2 of two of the three protomers at the apex of the envelope trimer.[58] All 96ZM651 alanine mutants contained in gp120 were screened with the CSAM procedure. In total, 26 mutants showed significantly altered binding towards PG9 (**Figure 27**). 12 of those were located in the inner domain of Env and 14 in the outer domain, with seven significant mutations in V1/V2 and two in V3. 25 of the statistically significant mutations led to a profound loss of PG9 binding, while mutant L111A depicted increased PG9 signals (**Figure 27**). Comparing the alanine scanning data to previously published crystallization and neutralization data,[117] most of the described contact residues between PG9 and Env could be identified by reduced PG9 binding with fold changes of 0.5 to 0.6 relative to 96ZM651 WT (**Figure 28A**). As PG9 bound weakly to 96ZM651 gp145, with MFIs of only about two- to threefold above background (**Table 4**), these fold changes already represent a prominent loss of antibody binding, although the weak signal intensities meant that only six of the described contact residues reached statistical significance. Of the described core epitope (N156, <u>N160</u>, <u>K168</u>, <u>K169</u>, Q170, <u>K171</u>, where underlined positions are especially important),[58] mutants N160A, K168A and K169A depicted a statistically significant decrease in PG9 binding with fold changes between 0.44 and 0.59, while K171A showed a slightly enhanced, though not significant, gain of PG9 signal intensity. Mutant N156A also exhibited a fold change of 0.59. Although not statistically significant, this also indicates the importance of this N-glycosylation site for PG9 binding. Q170A still depicted a fold change of 0.88, however, this position is only described as a contact residue in one of the two available co-crystallizations of PG9 with different V1/V2 peptides.[117] In general, most of the positions with a prominent loss of PG9 binding in the CSAM method were located at the trimer apex (**Figure 28B** and **C**), validating the mapping procedure for trimer-preferring[127] mAb PG9.

Most interestingly, L111A resulted in increased PG9 signal intensities, with a fold change of 2.2 compared to WT-like PG9 binding (**Figure 27C** and **D**). Position L111 is not located at the trimer apex and can therefore not be a direct interaction partner of PG9. Instead, the gp120 inner domain residue[128] faces the interaction surface of the different

trimer protomers (**Figure 28C**). L111A was previously described to strongly decrease aberrant dimerization of YU2[128,129] (clade B) and BG505[130] (clade A) gp120, thus facilitating gp120 monomer production. Furthermore, gp120 L111A mutants showed decreased affinity to sCD4 and the non-neutralizing antibodies targeting the CD4-induced Env conformation. Notably, this could be verified in the alanine scans of sCD4 and 17b, where L111A resulted in statistically significant reduction of binding with fold changes of 0.51 and 0.19, respectively (**Appendix Table 4**).
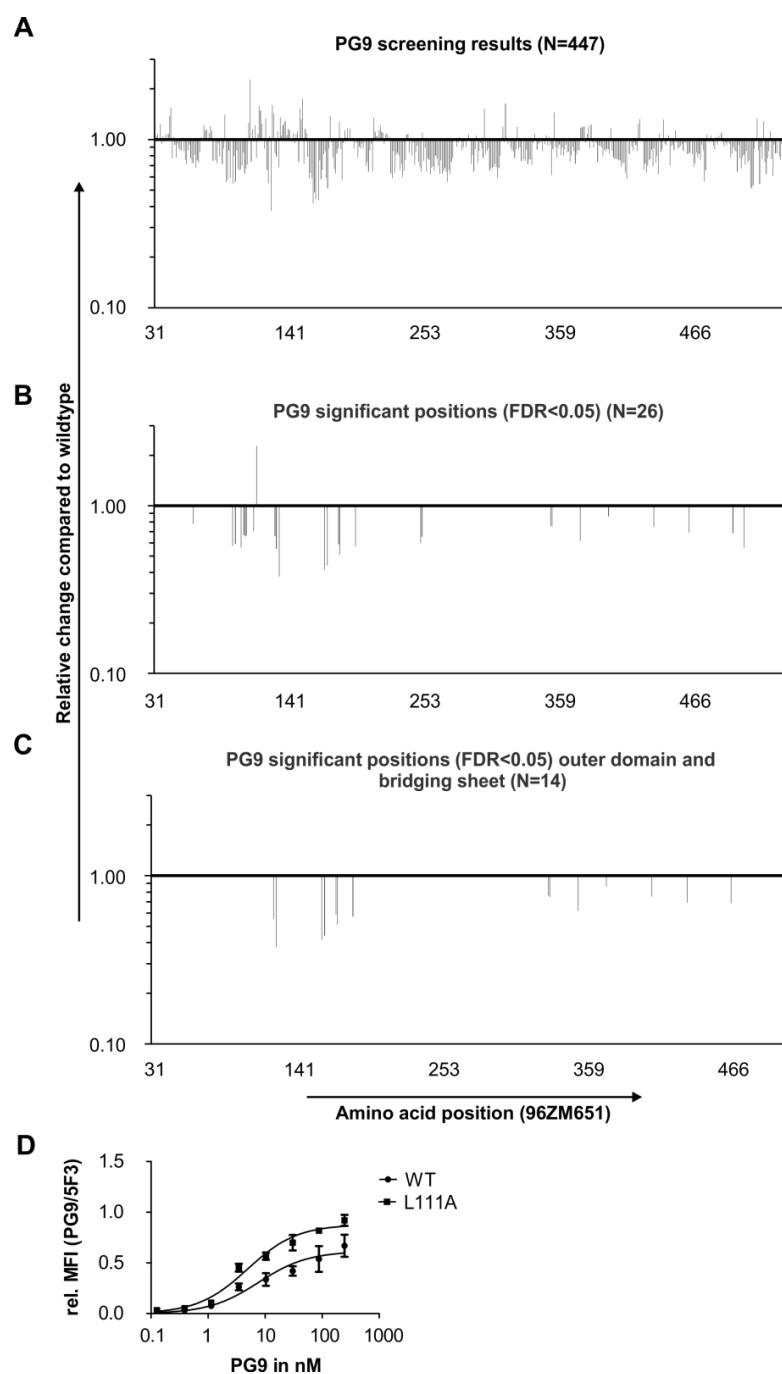
Figure 27: A: Alanine scan of PG9. 5F3 normalized fold changes of PG9 binding relative to 96ZM651 WT (WT≅1) are shown for every gp120 mutant of the alanine scanning library (n=6). The diagrams show **A:** all positions screened with the CSAM method (N=477), **B:** all positions which showed statistically significant fold changes compared to the WT (FDR<0.5, N=26) and **C:** all positions which reached significance and are located in the outer domain of Env and the bridging sheet (N=14) and are thus accessible from the surface. Statistical significance was determined in R using a two-sided one sample t-test against 1, while controlling the false discovery rate (FDR) to be smaller than 5 % to adjust for multiple testing. **D:** FACS titration of 96ZM651 gp145 WT and L111A with PG9. One of two biological replicates is shown, as varying MFI values did not allow pooling of the data. The second dataset, which depicted the same result, is shown in Appendix Figure 1.

A

Residue	FC	Structural effect
V127A	0.55*	disruption of V1V2 hydrophobic core
N156A	0.59	glycan-156 removal
C157A	NA	contact site in ZM109
S158A	0.42 *	glycan-156 removal, contact site in Cap45 and ZM109
F159A	0.49	disruption of V1V2 hydrophobic core
N160A	0.44 *	glycan -160 removal, side chain hydrogen bond contact site in Cap45 and ZM109
T162A	0.44	glycan-160 removal, contact site in Cap45 and ZM109
T163A	0.75	contact site in Cap45
E164A	0.58	contact site in Cap45
L165A	0.70	main chain hydrogen bond contact site in Cap45
K166A	0.82	contact site in Cap45 and ZM109
D167A	0.85	main chain hydrogen bond contact site in Cap45 and ZM109
K168A	0.59 *	side chain hydrogen bond contact site in Cap45 and ZM109
K169A	0.51 *	main chain hydrogen bond contact site in Cap45, main chain and side chain hydrogen bond contact site in ZM109
Q170A	0.88	side chain hydrogen bond contact site in Cap45
K171A	1.38	main chain hydrogen bond contact site in Cap45, main chain and side chain hydrogen bond contact site in ZM109
V172	0.96	contact site in Cap45 and ZM109
Y173A	0.78	side chain hydrogen bond contact site in ZM109, destabilization of glycan-156 position
F176A	0.63	disruption of V1V2 hydrophobic core
I181A	0.57 *	unknown as residues not seen in structures

0   1   2

Figure 28: Structural analysis of the FACS-based PG9 alanine scan. A: Binding data of the alanine scan for published contact residues of Env (HXB2 numbering) to PG9. The core epitope of PG9 as described in [58] is underlined and amino acids of special importance within the core epitope are bolded. 5F3 normalized fold changes (FC) of PG9 binding relative to 96ZM651 WT are given and statistically significant results (FDR<0.05) are asterisked. Contact residues and structural effects of mutating the respective positions were determined by co-crystallization of PG9 with V1/V2 peptides of isolates Cap45 and ZM109, or in neutralization assays.[117] **B:** Top view and **C:** Side view of the Env trimer (BG505 SOSIP.664, PDB ID: 4ZMJ). Amino acid mutations resulting in a fold change <0.6 and those of statistical significance (FDR<0.05) in the alanine scan are shown in red, the statistically significant gain of binding mutation in dark green (**C**, middle view) and the core epitope of PG9[58] in yellow. The overlap between the core epitope and the alanine scan is depicted in orange. The right side of **B** and all representations in **C** show one protomer of the gp140 trimer as a surface model, while the other protomers are illustrated as ribbon diagrams.

4.2.5.2. Permutation of significant positions from the PG9 alanine scan

PG9 recognizes the envelope trimer asymmetrically by interacting with variable loops V1 and V2 from two of the three trimer protomers.[127] Hence, significant loss of PG9 binding mutations located in V1/V2 as well as the gain of signal mutation L111A were chosen for permutation. For the loss of binding positions, this generally led to severely impaired PG9 recognition for almost all natural amino acids (**Figure 29B-K**). However, S158T resulted in a significant increase in PG9 signal intensity (**Figure 29D**). This serine to threonine mutation exchanges the wildtype sequence N-X-S to N-X-T. Both represent sequons for N-linked glycosylation, but N-X-T motives are thought to be more frequently and more extensively used,[131] supposedly because of their higher affinity to the oligosaccharyltransferase.[132] As the glycan at N156 is part of the core epitope of PG9,[58] S158T might therefore increase PG9 binding by influencing N-linked glycosylation at this site. However, mutant S158T depicted severely reduced expression levels (data not shown) and was not transferable to other isolates (**Appendix Table 5**). Hence, this mutation was not further pursued. Permutation at position L111, where the alanine scan detected a positive influence of L111A on PG9 binding, did not result in additional PG9 gain of signal mutants (**Figure 29A**).

To test whether these 11 alanine mutations have the same impact on other isolates, they were transferred to BG505 (clade A) and 16055 (clade C), which both served as a basis for SOSIP and NFL variants.[49,50,82] Most loss of PG9 binding mutations identified in the 96ZM651 alanine scan also severely impaired PG9 signals for both isolates. Only E164A, which depicted a fold change of 0.62 for 96ZM651, resulted approximately in wildtype-like binding levels in both BG505 and 16055. Notably, mutant L111A also showed significantly increased signal intensities for BG505 and 16055, with fold changes of 1.64 and 1.48, respectively (**Figure 29L**).
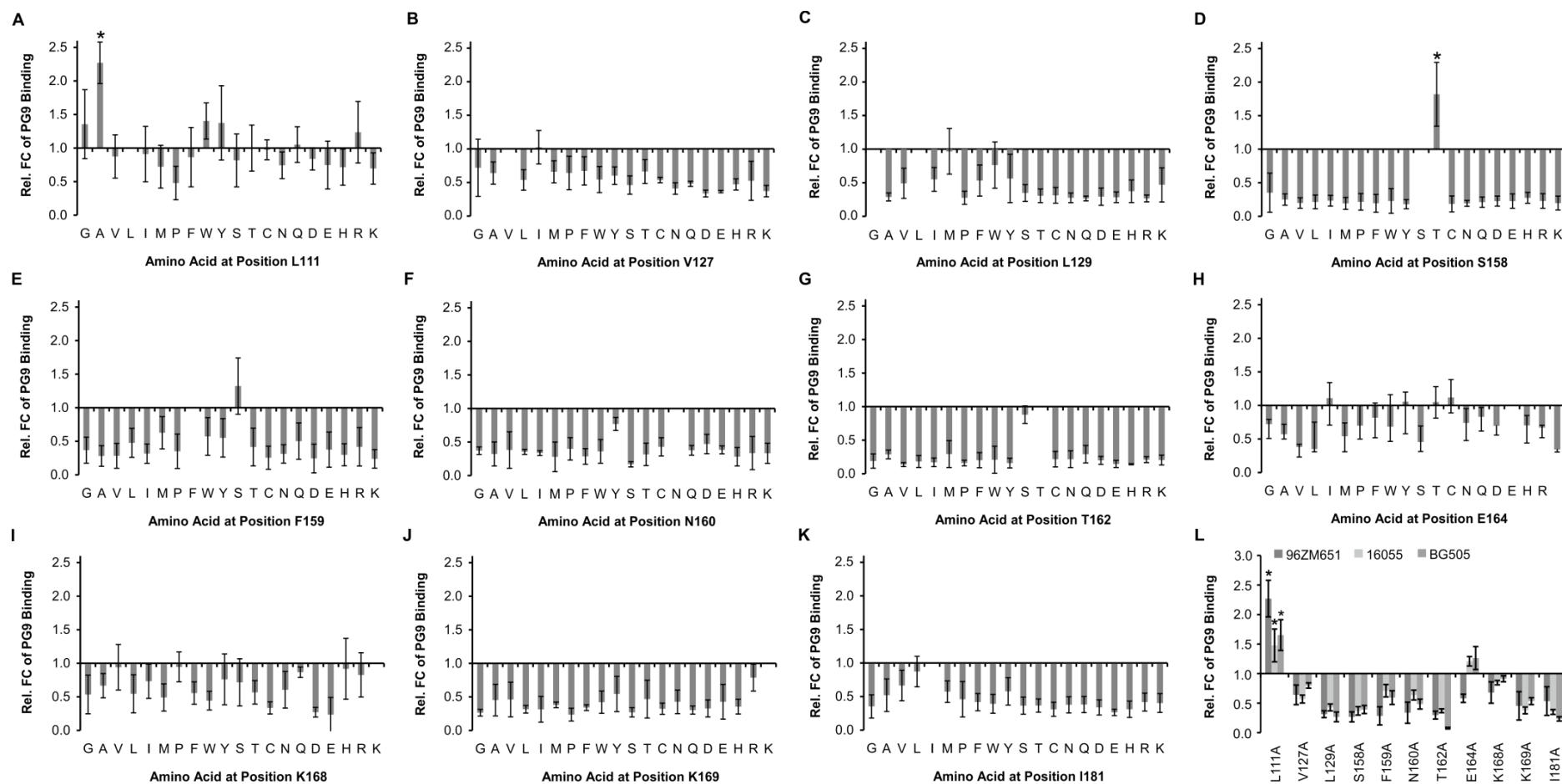
Figure 29: Permutation of significant positions from the PG9 alanine scan and transfer to other isolates. A to **K:** Position L111 (PG9 alanine scan gain of signal) as well as significant loss of signal mutations located in V1/V2 were permutated in 96ZM651 gp145 and mutants were analyzed for altered PG9 binding by flow cytometry (n=4-6). 5F3 normalized fold changes are indicated (WT≅1). Only S158T resulted in increased PG9 binding, while most other mutants also displayed a loss of PG9 signal. **L:** Transfer of significant loss of signal alanine mutations located in V1/V2 and the gain of signal mutation L111A to isolates 16055 (clade C) and BG505 (clade A). All mutations except E164A proofed to be transferable. Data were generated during the master thesis of Krystina Beer under my experimental supervision. Statistical significance was determined in R using a two-sided one sample t-test against 1, while adjusting the false discovery rate (FDR) to be smaller than 5 % to correct for multiple testing.

4.2.5.3.   Further characterization of mutant L111A

4.2.5.3.1. Impact of L111A on soluble gp140 trimer formation

The increased PG9 signal intensity observed for mutant L111A via flow cytometry could be caused by a direct increase in bnAb affinity. As PG9 is a trimer-preferring, quaternary structure-dependent antibody,[127] the increased signal may, however, also be the result of a higher trimer proportion on the cell surface. Although 5F3 normalizes for envelope surface presentation levels, it binds to a linear epitope[46] and thus presumably recognizes Env monomers, dimers and trimers alike. A shift in this distribution towards a higher trimer ratio would therefore not be easily visible in this flow cytometry-based assay format. As L111A is located in the inner part of the trimer (**Figure 28C**) and thus not part of the PG9 epitope, the gain of signal effect would be most likely caused by an influence of L111A on trimer structure. To further analyze this, soluble gp140 envelope trimers lacking the cytoplasmic tail as well as the transmembrane domain were expressed in HEK293-F suspension cells and purified from the supernatant using lectin affinity chromatography (**Figure 30A**) and subsequent size exclusion to reach trimer homogeneity (**Figure 30B**). Interestingly, insertion of L111A led to a profound reduction of monomeric gp140 (**Figure 30B** and **C**), indicating facilitated trimer formation for this mutant.
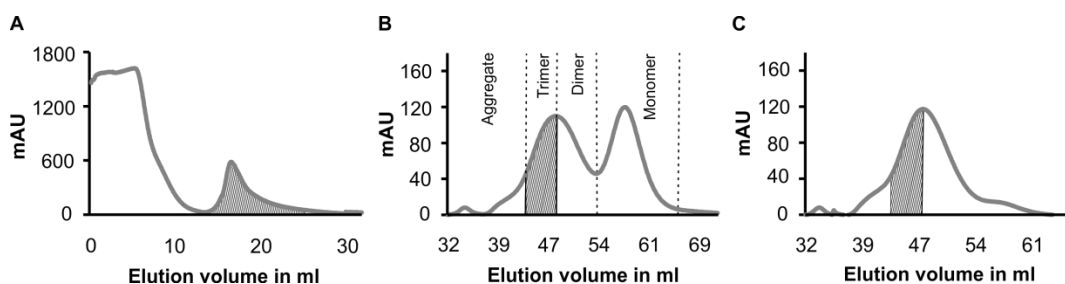
Figure 30: Exemplary elution profiles of 96ZM651 gp140 purifications. Soluble envelope gp140 trimers were purified from the supernatant of HEK293-F suspension cells by lectin affinity chromatography and subsequent size exclusion to selectively purify Env trimers. mAU: absorption units at 280 nm. **A:** elution profile of the *Galanthus nivalis* lectin chromatography. After washout of other proteins in the supernatant, Env was eluted with 1 M α-D-mannopyranoside, resuling in a single Env peak (highlighted in grey). **B:** Size exclusion chromatography elution profile of 96ZM651 gp140 WT. Env elutes as an aggregate shoulder, overlapping trimer (highlighted in gray) and dimer fractions and a monomer peak. **C:** Size exclusion chromatography elution profile of 96ZM651 gp140 L111A. The trimer fraction is highlighted in gray. Mutation L111A leads to a markedly reduced monomer peak.

4.2.5.3.2. Impact of L111A on Env thermostability

96ZM651 gp140 trimer fractions of wildtype and mutant L111A were subjected to differential scanning calorimetry (DSC). Both soluble trimers exhibited melting temperatures of approximately 60 °C, however, melting curves markedly differed between WT and mutant. Although the wildtype trimer was administered at approximately twice the concentration compared to L111A, the mutant depicted a more precise melting curve with considerably higher signal intensities (**Figure 31**). This is supposedly caused by a more homogeneous folding of mutant L111A compared to WT gp140, resulting in a more synchronous unfolding process for the variant and hence a more homogeneous melting curve. Measurements repeated with independent protein purifications and equal trimer concentrations yielded similar curves, excluding any influence of the purification process on these results (**Appendix Figure 4**).
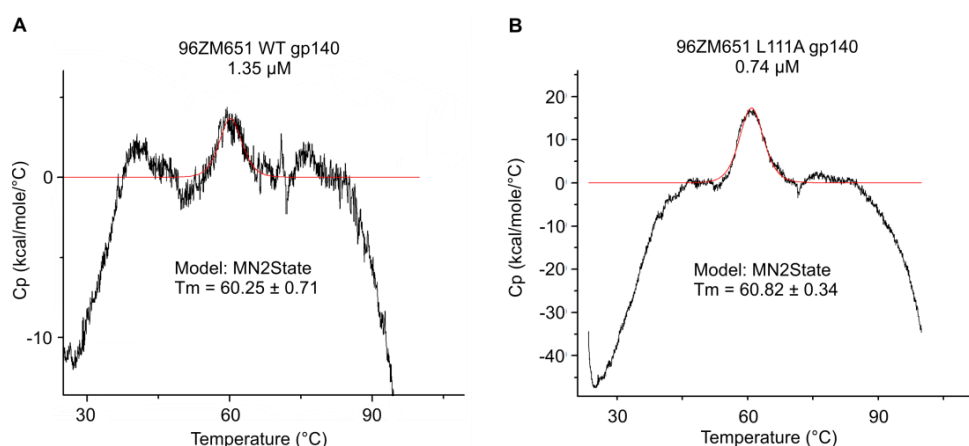
Figure 31: DSC analysis of 96ZM651 gp140 wildtype and mutant L111A. Trimer fractions of Env gp140 WT and L111A were purified by lectin affinity chromatography and subsequent size exclusion and analyzed by DSC. For 96ZM651 WT (**A**), a higher concentration (1.35 µM) was administered, but still only weak signal intensity and a very inhomogeneous peak was obtained. In the case of mutant L111A (**B**), a lower concentration (0.74 µM) already resulted in higher signal intensity and a more homogeneous melting curve, suggesting improved trimer folding. L111A did not alter Env thermostability (Tm: melting temperature). Red lines represent the fitted data used for Tm calculation.

4.2.5.3.3. Antigenicity of 96ZM651 L111A

Mutation L111A increased PG9 affinity to soluble 96ZM651 from a K_D value of 62.0 to 39.0 nM (**Tables 10** and **13**), further indicating improved Env folding of this variant. Moreover, comparing 96ZM651 WT and L111A for both membrane-bound gp145 via flow cytometry as well as for soluble gp140 by ELISA, mutant L111A showed preserved binding to bnAbs targeting different epitopes (**Table 9**). In addition to the positive influence on PG9 binding, improved binding was also observed for VRC01 (CD4BS), which was, although not statistically significant, already visible in the VRC01 alanine scan (**Appendix Table 3**). In addition, L111A resulted in slightly increased signal intensities for bnAb 2G12, which binds to a conserved mannose-patch on gp120.[45] Interestingly, in agreement with data for BG505 gp120,[129] L111A also decreased sCD4 binding to gp145 and severely impaired recognition of non-neutralizing mAb 17b for both membrane-bound and soluble envelope (**Table 9**).

Table 9: Impact of L111A on different epitopes on membrane-bound and soluble 96ZM651 envelope. gp145 L111A was analyzed by flow cytometry (n=3), while gp140 trimer fractions were analyzed by ELISA (n=3), with utilized antibody concentrations determined from Appendix Figure 3. Fold changes are given relative to 96ZM651 WT (WT≙1). For gp145, values were normalized with 5F3. L111A increased signals of PG9, VRC01 and 2G12 while other bnAb epitopes were preserved. L111A also reduced sCD4 (gp145) and 17b binding (gp145 and gp140). Data for gp145 were obtained during the master thesis of Matthias Glögl under my experimental supervision.

		broadly neutralizing							non or weakly neutralizing		
	Mutant	MPER		Glycan	V3		CD4BS	Apex	CD4i	CD4BS	
		4E10	10E8	2G12	HGN194	PGT121	VRC01	PG9	17b	b6	sCD4
membrane-bound	L111A gp145	1.00 ± 0.11	0.88 ± 0.02	1.65 ± 0.28	0.81 ± 0.11	1.08 ± 0.17	1.43 ± 0.33	1.83 ± 0.34	0.20 ± 0.02	0.91 ± 0.18	0.51 ± 0.19
soluble	L111A gp140	0.93 ± 0.12	1.06 ± 0.11	1.28 ± 0.02	0.98 ± 0.01	1.13 ± 0.02	1.48 ± 0.11	1.33 ± 0.02	0.53 ± 0.03	1.11 ± 0.01	1.08 ± 0.01

0   1   5

4.2.5.3.4. Transfer of L111A to other isolates

To analyze its effect on other isolates, L111A was transferred to membrane-bound gp145 and soluble gp140 of BG505 (clade A), HXB2 (clade B), JRFL (clade B) and 16055 (clade C), as well as BG505 SOSIP and 16055 SOSIP. Analysis of gp145 variants by flow cytometry resulted in increased PG9 signals for BG505 and 16055 SOSIP, although 96ZM651 still displayed the highest fold change. BG505 SOSIP, HXB2 and 16055 showed approximately wildtype-like binding, while JRFL exhibited a slightly decreased fold change. Similarly, apart from 96ZM651, increased fold changes of VRC01 signal intensity were observed for BG505 and BG505 SOSIP, as well as 16055 SOSIP. JRFL showed unchanged VRC01 binding, while signals for HXB2 and 16055 were slightly reduced (**Table 10**, left side). In contrast, L111A in soluble gp140 resulted in heightened affinity to PG9 and VRC01 for most isolates, with fold changes of approximately two. Only 16055 SOSIP gp140 L111A showed approximately WT-like affinity to both bnAbs, while HXB2 exhibited decreased affinity to PG9 (**Table 10**, right side). In sum, the impact of L111A on antigenicity proved to be transferable to most isolates tested. However, PG9 fold changes for gp145 and gp140 variants were not always in agreement (e.g. for JRFL, which did not show increased signal intensities for gp145 in the flow cytometry assay but exhibited enhanced affinities to both PG9 and VRC01 in the case of soluble gp140). This supports the hypothesis that different effects like shifted monomer to trimer ratios on the cell surface and altered Env folding play a role for L111A mutants, impairing the correlation between PG9 MFIs and its affinity to the respective variant. Notably, in accordance with T278 mutants (**Table 8**), magnitudes of VRC01 and PG9 affinity varied between the different isolates, with 96ZM651 exhibiting particularly low affinity to both bnAbs.

Table 10: Transfer of L111A to other isolates. gp145 L111A variants were analyzed for altered PG9 and VRC01 binding by flow cytometry. 5F3 normalized fold changes (FC) are indicated relative to the respective WT (n=3). gp140 L111A trimers were analyzed via ELISA titration (Appendix Figure 2, n=2). L111A increased signal intensities (gp145) as well as affinity of PG9 and VRC01 (gp140) for most tested constructs. Data (except 96ZM651) were generated during the master thesis of Christina Schmalzl under my experimental supervision.

Clade	Isolate	Variant	cell surface gp145		soluble gp140			
			FC of PG9 binding to membrane-bound gp145	FC of VRC01 binding to membrane-bound 96ZM651 gp145	K_D of PG9 binding to soluble gp140 in nM	FC of PG9 binding to soluble gp140	K_D of VRC01 binding to soluble gp140 in nM	FC of VRC01 binding to soluble gp140 gp140
A	BG505	WT	1	1	6.6 ± 1.4	1.8	4.3 ± 0.4	1.8
		L111A	1.78 ± 0.15	2.02 ± 0.58	3.6 ± 1.0		2.5 ± 0.4	
A	BG505 SOSIP	WT	1	1	35.3 ± 8.5	2.1	12.3 ± 2.2	1.6
		L111A	1.22 ± 0.06	1.30 ± 0.14	16.6 ± 1.3		7.6 ± 0.7	
B	HXB2	WT	1	1	23.0 ± 3.0	0.3	11.0 ± 0.8	1.0
		L111A	0.90 ± 0.09	0.71 ± 0.11	69.7 ± 27.5		11.4 ± 1.5	
B	JRFL	WT	1	1	32.6 ± 13.3	2.7	6.8 ± 0.5	1.6
		L111A	0.69 ± 0.07	0.96 ± 0.09	12.3 ± 3.1		4.3 ± 0.4	
C	16055	WT	1	1	6.8 ± 1.1	1.8	12.2 ± 0.1	2.2
		L111A	1.10 ± 0.10	0.62 ± 0.07	3.7 ± 1.0		5.6 ± 0.4	
C	16055 SOSIP	WT	1	1	6.0 ± 1.3	0.9	13.5 ± 1.0	1.3
		L111A	1.71 ± 0.09	1.43 ± 0.24	6.4 ± 1.8		10.6 ± 1.5	
C	96ZM651	WT	1	1	62.0 ± 17.3	1.6	38.7 ± 2.7	1.7
		L111A	2.27 ± 0.14	1.54 ± 0.21	39.0 ± 8.1		22.4 ± 4.2	
			0 1 highest FC	0 1 highest FC	highest lowest K_D K_D	0 1 highest FC	highest lowest K_D K_D	0 1 highest FC

Notably, L111A also resulted in altered SEC elution profiles for all soluble trimers. Similar to 96ZM651, the gp140 monomer peak was markedly reduced for HXB2, JRFL and 16055, while trimer yield was distinctly increased for 16055 SOSIP. In the case of BG505, the monomer peak was shifted to a higher elution volume, indicating a tighter folding of the protein, while a newly arising shoulder for BG505 SOSIP suggests a better separation of trimer, dimer and monomer fractions. However, for BG505 SOSIP, L111A also negatively influenced Env expression (**Figure 32**). In accordance with results for 96ZM651, L111A did not alter thermostability in any of the variants tested (**Table 11**).

In conclusion, L111A resulted in facilitated gp140 trimer formation as well as increased PG9 and VRC01 affinity for most of the different isolates and SOSIP constructs tested.
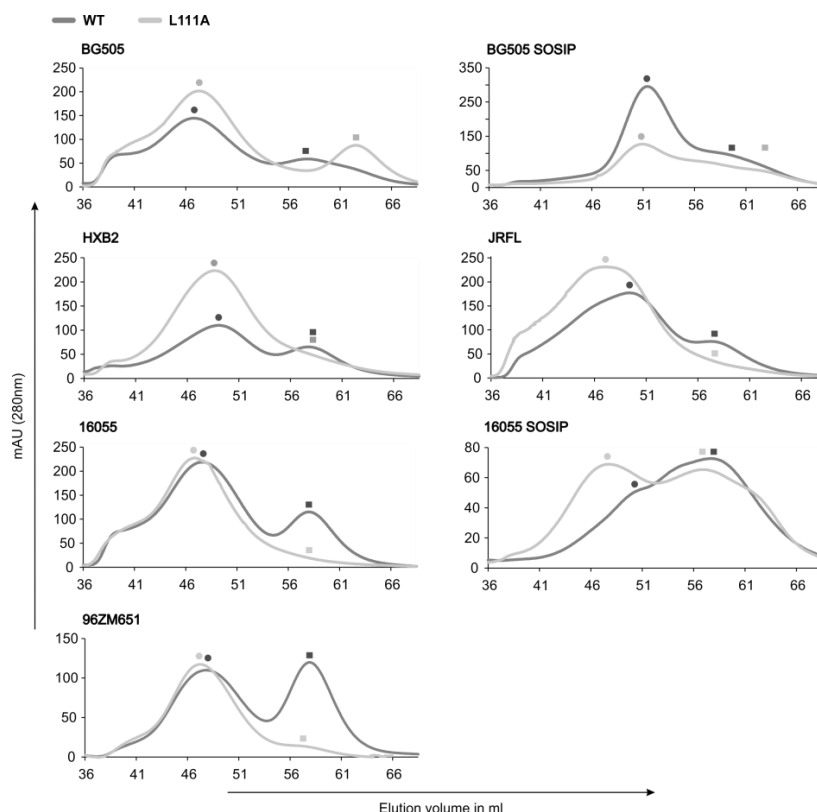
Figure 32: Size exclusion elution profiles for gp140 WT and L111A envelopes from different isolates. gp140 trimers were purified from the supernatant of HEK293-F suspension cells with lectin affinity chromatography and subsequent size exclusion (SEC). SEC elution profiles are shown with WT proteins in dark grey and L111A variants in light gray. Trimer fractions are marked by circles, monomer peaks with squares. L111A altered the elution profile of all tested constructs. Data for all isolates except 96ZM651 were generated during the master thesis of Christina Schmalzl under my experimental supervision.

Clade	Isolate	T_M WT	T_M L111A
A	BG505	59.6 ± 7.1	56.9 ± 0.1
A	BG505 SOSIP	65.7 ± 0.8	66.8 ± 0.2
B	HXB2	58.8 ± 1.8	58.3 ± 0.2
B	JRFL	57.5 ± NA	56.6 ± NA
C	16055	66.7 ± 15.7	68.2 ± 16.2
C	16055 SOSIP	NA	NA
C	96ZM651	58.4 ± 1.8	58.9 ± 0.7

Table 11 : NanoDSF analysis of gp140 WT and L111A trimers of different isolates. (n=2) Melting temperatures (T_M) were not influenced by L111A.

4.2.6. Combination of significant mutations from the sCD4, VRC01 and PG9 alanine scans

In summary, the performed alanine scans and subsequent permutation identified several mutations which positively influenced Env antigenicity. W427A showed profoundly decreased binding of sCD4, which would supposedly prohibit the opening of trimeric Env immunogens *in vivo* and thus decrease non-neutralizing antibody titers.[79] Notably, this is further supported by the markedly reduced binding of non-neutralizing CD4i mAb 17b (**Table 5**). In addition, mutants T278A and H increased affinity of CD4BS-targeting bnAb VRC01, with T278H being slightly superior for most isolates tested and supposedly having a positive influence on Env folding (**Tables 8** and **13**). Furthermore, L111A enhanced trimer formation and apparently improved Env folding (**Figures 30** to **32**). This resulted in increased affinities of both PG9 and VRC01 to most tested isolates (**Table 10**), while simultaneously reducing 17b binding (**Table 9**). Notably, none of these mutations significantly impaired other bnAb epitopes (**Tables 5**, **7** and **9**). Hence, 96ZM651 gp145 combination mutants of L111A, T278A or H and W427A were generated to combine these favorable phenotypes into a single Env immunogen (**Table 12**, upper panel). While other bnAb epitopes could be preserved, double mutants L111A/T278A and L111A/T278H both exhibited enhanced PG9 and VRC01 binding. In accordance with T278 single mutants, L111A/T278H thereby depicted the superior effect for both bnAbs. Notably, L111A strongly reduced 17b binding in both double mutants. However, L111A/T278H showed slightly increased sCD4 recognition. In agreement with previous results (**Table 5**), inclusion of W427A resulted in decreased sCD4 binding. While sCD4 recognition was almost completely abolished for L111A/T278A, the slight enhancement of sCD4 binding observed for T278H counteracted this effect, resulting in a fold change of 0.59 for L111A/T278H/W427A. However, in addition to impairing sCD4 recognition, W427A also slightly decreased the gain of binding effects for PG9 and VRC01 (**Table 12**). Therefore, as a potentially beneficial effect of loss of CD4 binding on immunogenicity would only be visible in a more complex, human CD4 transgenic or humanized animal model, only double mutants L111A/T278A and L111A/T278H and one selected triple mutant (L111A/T278A/W427A) were analyzed as soluble trimers. As already observed for gp140 L111A, all purified double and triple mutants also showed improved trimer formation (**Figure 30**), rendering the monomer fractions even smaller than for the trimer-stabilized BG505 SOSIP (**Figure 33**).
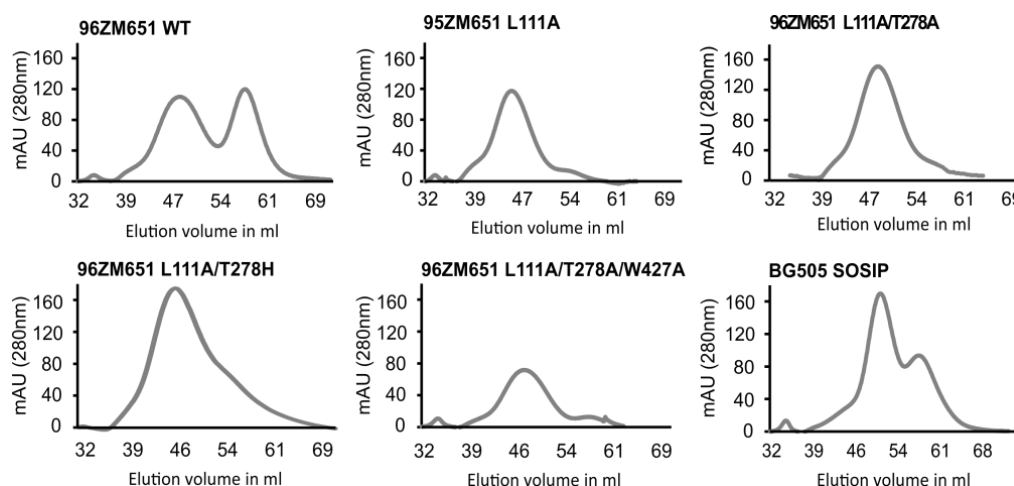
Figure 33: Exemplary size exclusion elution profiles for 96ZM651 gp140 WT, L111A, and selected combination mutants. L111A reduced monomer peaks in all variants.

Although fold changes for PG9 and VRC01 were not as enhanced as for gp145, antigenicity profiles of membrane-bound variants were well transferable to soluble 96ZM651 gp140 double and triple mutants (**Table 12**, lower panel). Notably, the slightly enhanced sCD4 binding of mutant T278H as observed for gp145 was thereby not confirmed for soluble trimers. In agreement with membrane-bound gp145, 96ZM651 L111A/T278H depicted the most pronounced gain of binding effects for both VRC01 and PG9. ELISA titration of 96ZM651 single, double and triple mutants with both bnAbs confirmed these results, with gp140 T278H and L111A/T278H depicting approximately threefold increased affinities to VRC01 (**Table 13**). Most interestingly, however, the supposed structural effects of L111A and T278H seemed to act additive regarding PG9 affinity, resulting in a K_D value of only 9.3 nM for L111A/T278H double mutant and hence an improvement of nearly sevenfold compared to 96ZM651 WT (**Table 13**). This could be independently confirmed during the master thesis of Helene Hierl (**Table 15**). Thermostability, which was shown to correlate with serum neutralization activity after vaccination,[50] was preserved for all tested 96ZM651 gp140 trimers (**Table 13**).

In conclusion, both membrane-bound and soluble envelope immunogens with increased affinity to bnAbs VRC01 and PG9 could be generated, while bnAb binding to other epitopes was preserved. Simultaneously, recognition of both sCD4 and non-neutralizing CD4i mAb 17b was reduced, further supporting a neutralizing antibody response *in vivo*.

In summary, double mutant L111A/T278H thereby exhibited the most promising antigenicity profile.
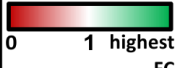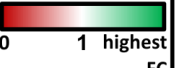
Table 12: Influence of L111A, T278A/H and W427A combination on different epitopes on membrane-bound and soluble 96ZM651 envelope. gp145 variants were analyzed by flow cytometry (n=3), while soluble gp140 trimers were analyzed by ELISA (n=2), with applied antibody concentrations determined from Appendix Figure 3. Fold changes are indicated relative to 96ZM651 gp145 or gp140 WT, respectively. In the case of gp145, fold changes were normalized with 5F3. Antigenicity effects of the single mutants were observed to be combinable in both membrane-bound and soluble 96ZM651 envelope.

			broadly neutralizing							non or weakly neutralizing		
		Mutant	MPER		Glycan	V3		CD4 BS	Apex	CD4i	CD4BS	
			4E10	10E8	2G12	HGN194	PGT121	VRC01	PG9	17b	b6	sCD4
Membrane-bound gp145	Single Mutants	L111A	1.00 ± 0.11	0.88 ± 0.02	1.65 ± 0.28	0.81 ± 0.11	1.08 ± 0.17	1.43 ± 0.33	1.83 ± 0.34	0.20 ± 0.02	0.91 ± 0.18	0.51 ± 0.19
		T278A	0.96 ± 0.17	1.11 ± 0.13	1.32 ± 0.44	1.06 ± 0.12	1.10 ± 0.17	1.73 ± 0.51	1.20 ± 0.06	0.80 ± 0.27	0.93 ± 0.21	0.62 ± 0.27
		T278H	1.17 ± 0.15	1.07 ± 0.01	0.97 ± 0.06	1.10 ± 0.01	1.22 ± 0.03	5.32 ± 0.10	3.37 ± 0.26	1.35 ± 0.06	1.09 ± 0.01	1.43 ± 0.13
	Double Mutants	L111A T278A	0.94 ± 0.16	0.92 ± 0.09	1.62 ± 0.49	0.98 ± 0.17	1.11 ± 0.04	2.28 ± 0.55	2.01 ± 0.65	0.16 ± 0.07	0.61 ± 0.12	0.87 ± 0.20
		L111A T278H	1.07 ± 0.06	1.06 ± 0.17	0.87 ± 0.14	1.06 ± 0.18	1.16 ± 0.02	5.58 ± 0.44	2.37 ± 0.05	0.21 ± 0.02	0.37 ± 0.05	1.78 ± 0.04
	Triple Mutants	L111A T278A W427A	0.81 ± 0.11	0.87 ± 0.08	1.03 ± 0.30	0.88 ± 0.13	0.98 ± 0.14	1.36 ± 0.29	1.74 ± 0.32	0.11 ± 0.04	0.82 ± 0.30	0.03 ± 0.00
		L111A T278H W427A	1.94 ± 0.16	1.13 ± 0.05	0.80 ± 0.10	1.20 ± 0.13	1.19 ± 0.16	3.39 ± 0.57	2.35 ± 0.56	0.05 ± 0.01	1.45 ± 0.07	0.59 ± 0.09
			4E10	10E8	2G12	HGN194	PGT121	VRC01	PG9	17b	b6	sCD4
Soluble gp140	Single Mutants	L111A	0.93 ± 0.12	1.06 ± 0.11	1.28 ± 0.02	0.98 ± 0.01	1.13 ± 0.02	1.48 ± 0.11	1.33 ± 0.02	0.53 ± 0.03	1.11 ± 0.01	1.08 ± 0.01
		T278A	0.87 ± 0.06	0.86 ± 0.10	0.75 ± 0.03	0.89 ± 0.06	0.67 ± 0.00	1.29 ± 0.02	1.13 ± 0.21	0.81 ± 0.08	0.89 ± 0.03	0.80 ± 0.05
		T278H	1.04 ± 0.00	1.02 ± 0.01	0.87 ± 0.33	1.01 ± 0.02	1.14 ± 0.07	1.72 ± 0.60	1.65 ± 0.13	0.96 ± 0.08	1.04 ± 0.01	0.96 ± 0.02
	Double Mutants	L111A T278A	0.83 ± 0.17	0.82 ± 0.15	0.86 ± 0.04	0.83 ± 0.02	1.15 ± 0.10	2.61 ± 0.15	1.77 ± 0.08	0.20 ± 0.01	0.90 ± 0.09	0.97 ± 0.04
		L111A T278H	1.04 ± 0.09	1.04 ± 0.15	1.73 ± 0.41	1.02 ± 0.06	1.21 ± 0.12	2.14 ± 0.89	2.45 ± 0.62	0.49 ± 0.38	1.06 ± 0.04	1.01 ± 0.01
	Triple Mutants	L111A T278A W427A	1.01 ± 0.15	1.10 ± 0.12	1.02 ± 0.05	1.23 ± 0.06	1.00 ± 0.12	1.44 ± 0.06	1.54 ± NA	0.13 ± 0.00	0.97 ± 0.02	0.06 ± 0.00
		L111A T278H W427A	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

0 1 5

Table 13: Thermostability of 96ZM651 single, double and triple mutants and their affinity to PG9 and VRC01. A summary of 5F3 normalized fold changes to VRC01 and PG9 as determined by flow cytometry for gp145 (n=3), K_D values for gp140 trimers as determined by ELISA titration (Appendix Figure 2, n=2) and melting temperatures of soluble trimers (T_M) as determined by nanoDSF (n=2) is given.

	cell surface gp145		soluble gp140		
	fold change of bnAb binding compared to 96ZM651 gp145 WT		K_D in nM		T_M in °C
Variant	VRC01	PG9	VRC01	PG9	
WT	1.00	1.00	38.7 ± 2.7	62.0 ± 17.3	58.4 ± 1.8
L111A	1.43 ± 0.33	1.83 ± 0.34	22.4 ± 4.2	39.0 ± 8.1	58.9 ± 0.7
T278A	1.73 ± 0.51	1.20 ± 0.06	16.1 ± 0.6	53.9 ± 14.9	56.6 ± 0.1
T278H	5.32 ± 0.10	3.37 ± 0.26	11.4 ± 0.6	30.3 ± 5.9	NA
L111A/T278A	2.28 ± 0.55	2.01 ± 0.65	20.7 ± 7.1	25.7 ± 4.8	57.5 ± 0.1
L111A/T278H	5.58 ± 0.44	2.37 ± 0.05	12.9 ± 2.2	9.3 ± 1.8	61.1 ± 0.7
L111A/T278A/W427A	1.36 ± 0.29	1.74 ± 0.32	17.1 ± 7.3	44.8 ± 14.3	59.5 ± 0.8
L111A/T278H/W427A	3.39 ± 0.57	2.35 ± 0.56	NA	NA	NA
	0　　1　highest FC	0　　1　highest FC	highest K_D　lowest K_D	highest K_D　lowest K_D	

Next, L111A/T278A, L111A/T278H, L111A/T278A/W427A and L111A/T278H/W427A were applied to the previously applied gp145 panel of different isolates and SOSIP variants. In agreement with data collected for the single mutants (**Table 10**), the beneficial impact of L111A on PG9 binding was not visible for all isolates, presumably due to an interplay of different effects (e.g. altered trimerization, Env folding and affinity). Also in accordance with the single mutants (**Table 8**), VRC01 signals were increased for all double variants tested, with 16055 SOSIP displaying the most pronounced fold change and L111A/T278H at a slight advantage in all tested isolates. Similar to 96ZM651 combination mutants (**Table 12**), addition of W427A markedly reduced sCD4 binding for almost all triple variants, but also diminished the enhancement of VRC01 and PG9 binding (**Table 14**).

Table 14: Flow cytometry analysis of membrane-bound gp145 combination mutants from different isolates. 5F3 normalized fold changes relative to the respective wildtype are indicated for binding of PG9, VRC01 and sCD4 as determined by flow cytometry (n=3). While a PG9 gain of signal was not observed for most isolates, higher signal intensities were obtained for VRC01, with T278H-containing mutants at a slight advantage. Inclusion of W427A resulted in almost complete abolishment of sCD4 binding, but also decreased bnAb recognition.

	Clade	gp145	FC PG9		FC VRC01		FC sCD4	
			L111A/T278A	L111A/T278H	L111A/T278A	L111A/T278H	L111A/T278A	L111A/T278H
double mutants	A	BG505	1.01 ± 0.26	1.63 ± 0.31	1.28 ± 0.51	1.75 ± 0.35	0.67 ± 0.39	0.89 ± 0.26
	A	BG505 SOSIP	0.93 ± 0.13	1.21 ± 0.14	2.05 ± 0.67	2.86 ± 0.81	0.43 ± 0.13	0.54 ± 0.16
	B	HXB2	0.64 ± 0.12	0.55 ± 0.13	1.65 ± 0.53	1.63 ± 0.53	0.85 ± 0.34	0.72 ± 0.26
	B	JRFL	0.63 ± 0.20	0.45 ± 0.06	1.80 ± 0.54	1.88 ± 0.55	1.51 ± 0.42	1.35 ± 0.37
	C	16055	0.83 ± 0.18	0.89 ± 0.17	1.52 ± 0.26	1.72 ± 0.46	1.01 ± 0.39	0.94 ± 0.45
	C	16055 SOSIP	1.33 ± 0.25	1.46 ± 0.25	3.11 ± 0.39	3.21 ± 0.55	1.21 ± 0.24	1.15 ± 0.23
	C	96ZM651	1.89 ± 0.41	1.12 ± 0.14	2.28 ± 0.55	2.50 ± 0.21	0.87 ± 0.20	0.80 ± 0.09

	Clade	gp145	FC PG9		FC VRC01		FC sCD4	
			L111A/T278A/W427A	L111A/T278H/W427A	L111A/T278A/W427A	L111A/T278H/W427A	L111A/T278A/W427A	L111A/T278H/W427A
triple mutants	A	BG505	0.72 ± 0.49	0.90 ± 0.25	0.87 ± 0.16	1.25 ± 0.38	0.31 ± 0.10	0.41 ± 0.19
	A	BG505 SOSIP	0.45 ± 0.05	0.66 ± 0.10	1.62 ± 0.35	2.45 ± 0.63	0.06 ± 0.02	0.07 ± 0.02
	B	HXB2	0.93 ± 0.30	0.78 ± 0.17	1.56 ± 0.47	1.48 ± 0.40	0.42 ± 0.20	0.34 ± 0.13
	B	JRFL	0.84 ± 0.26	0.63 ± 0.23	1.52 ± 0.38	1.55 ± 0.42	1.21 ± 0.37	1.10 ± 0.41
	C	16055	0.56 ± 0.08	1.64 ± 0.22	1.36 ± 0.50	2.54 ± 0.49	0.06 ± 0.01	0.61 ± 0.37
	C	16055 SOSIP	0.92 ± 0.14	1.03 ± 0.16	1.76 ± 0.22	1.96 ± 0.20	0.06 ± 0.02	0.05 ± 0.01
	C	96ZM651	1.74 ± 0.32	0.80 ± 0.21	1.36 ± 0.29	1.46 ± 0.05	0.03 ± 0.00	0.03 ± 0.00

| | | | 0   1   3 | | 0   1   3 | | 0   1   3 | |

In summary, double mutant L111A/T278H showed the most promising binding to VRC01 and PG9 for both membrane-bound and soluble 96ZM651 (**Table 12**) as well as for the membrane-bound variants of all other tested isolates and SOSIP constructs (**Table 14**). Hence, soluble L111A/T278H gp140 trimers were purified for BG505, BG505 SOSIP, 16055 and 16055 SOSIP, which are currently prominent in the literature.[50,85,130] In agreement with previously obtained data for L111A single mutants (**Figure 32**), the favorable effect of L111A on Env trimer formation was also visible for 16055 and 16055 SOSIP, but not BG505 and BG505 SOSIP double mutants.
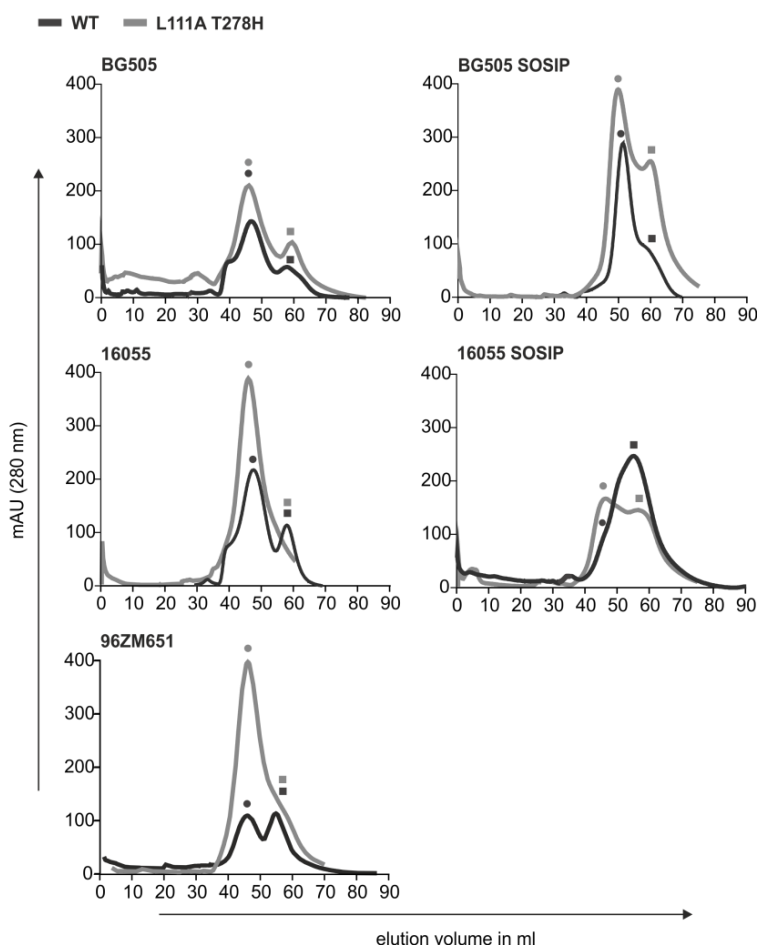
Figure 34: Exemplary size exclusion elution profiles of WT and L111A/T278H gp140 envelopes from different isolates. gp140 proteins of BG505 (clade A), 16055 (clade C), 96ZM651 (clade C) as well as BG505 SOSIP and 16055 SOSIP were purified from the supernatant of HEK293-F suspension cells with lectin affinity chromatography and subsequent size exclusion (SEC). SEC elution profiles of WT proteins are shown in black and of L111A/T278H variants in gray. All purifications were conducted from an equal volume of cell suspension, except for BG505 SOSIP, where the L111A/T278H mutant was purified from a threefold volume compared to BG505 WT. Trimer fractions are marked by circles, monomer peaks with squares. L111A improves trimer formation in the double mutants of 96ZM651, 16055 and 16055 SOSIP. Data were generated during the master thesis of Helene Hierl under my experimental supervision.

In accordance with 96ZM651 single mutants (**Tables 8** and **10**), L111A/T278H variants of all tested isolates showed decreased K_D values for VRC01 and PG9 (**Table 15**). Notably, however, magnitudes of K_D values again varied between the different isolates, with 96ZM651 exhibiting markedly lower affinity to both bnAbs.

Clade	Isolate	Variant	K_D to PG9 in nM	FC of PG9 affinity	K_D to VRC01 in nM	FC of VRC01 affinity
A	BG505	WT	0.6 ± 0.8	2.0	0.2 ± 0.2	5.0
		L111A/T278H	0.3 ± 2.1		0.04 ± 0.01	
A	BG505 SOSIP	WT	4.8 ± 2.9	2.0	0.8 ± 0.4	4.0
		L111A/T278H	2.4 ± 1.4		0.2 ± 0.1	
C	16055	WT	0.5 ± 0.3	2.5	0.6 ± 0.4	6.0
		L111A/T278H	0.2 ± 0.1		0.1 ± 0.0	
C	16055 SOSIP	WT	0.5 ± 0.3	1.7	0.5 ± 0.3	5.0
		L111A/T278H	0.3 ± 0.2		0.1 ± 0.1	
C	96ZM651	WT	49.7 ± 7.9	6.3	4.7 ± 1.2	47.0
		L111A/T278H	7.9 ± 1.3		0.1 ± 0.0	
			highest K_D lowest K_D	0 1 highest FC	highest K_D lowest K_D	0 1 highest FC

Table 15: Affinities of L111A/T278H gp140 trimers of different isolates to VRC01 and PG9. BG505, BG505 SOSIP, 16055, 16055 SOSIP and 96ZM651 trimers were titrated with PG9 and VRC01 via ELISA (n=2-3). All mutants showed increased bnAb affinity. Data were generated by Helene Hierl under my experimental supervision.

4.2.7. Stabilization of L111A/T278H envelope trimers by chemical cross-linking

Recently published findings suggest that cross-linking of envelope trimers supports well-ordered trimer integrity, thus reducing the exposure of non-neutralizing epitopes located in the inner part of the trimer.[78,79] In addition, increased thermostability of cross-linked envelope antigens correlates with the neutralization capacity of induced antibody responses.[50] Hence, L111A/T278H gp140 trimers of 96ZM651, BG505, BG505 SOSIP, 16055 and 16055 SOSIP were subjected to cross-linking by glutaraldehyde (GLA) and EDC/NHS (1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride/N-hydroxysuccinimide). The homo-bifunctional aldehyde GLA links amine groups, with lysine residues most commonly affected. Thereby, a five-residue carbon chain is inserted between the two cross-linked groups. In contrast, the hetero-bifunctional EDC reaction, which is catalyzed by NHS, cross-links an amino- to a carboxyl-group without inserting additional atoms.[79] All experiments regarding cross-linking were performed by Helene Hierl under my experimental supervision.

To determine optimal cross-linking rates for each isolate, all variants were treated with different molar ratios of Env to GLA or EDC/NHS. The success of the cross-linking reactions was visualized by SDS-PAGE and further analyzed by ELISA using (i) VRC01 and PG9 to assess the maintenance of bnAb epitopes and L111A/T278H gain of binding effects and (ii) non-neutralizing CD4i antibody 17b to evaluate well-folded trimer integrity. In accordance with literature data,[79,50] two different ELISA formats were performed: a capture ELISA, in which Env trimers were captured by *Galanthus nivalis* lectin to preserve trimer structure, as well as a direct ELISA, in which gp140 trimers were directly coated onto the plate, resulting in hydrophobic interactions and hence mildly denaturating conditions. For both settings, area under the curve values (AUC)[78,79,92] were determined as a means to compare antigenicity profiles of the differently cross-linked trimers. While we wanted to preserve at least 75 % of VRC01 and PG9 recognition, 17b binding should be reduced to less than 50 % compared to unmodified trimers.

For GLA, all cross-linking concentrations which reduced 17b signals to fewer than 50 % also severely impaired PG9 recognition (**Figure 35**). In contrast, EDC/NHS cross-linking could preserve both VRC01 and PG9 binding while simultaneously reducing 17b recognition. While 160 000 mol EDC/NHS per mol envelope met the above defined

criteria for 16055, 16055 SOSIP, BG505 and BG505 SOSIP L111A/T278H variants, 35 000x EDC/NHS to Env qualified for 96ZM651 (**Figure 36**).
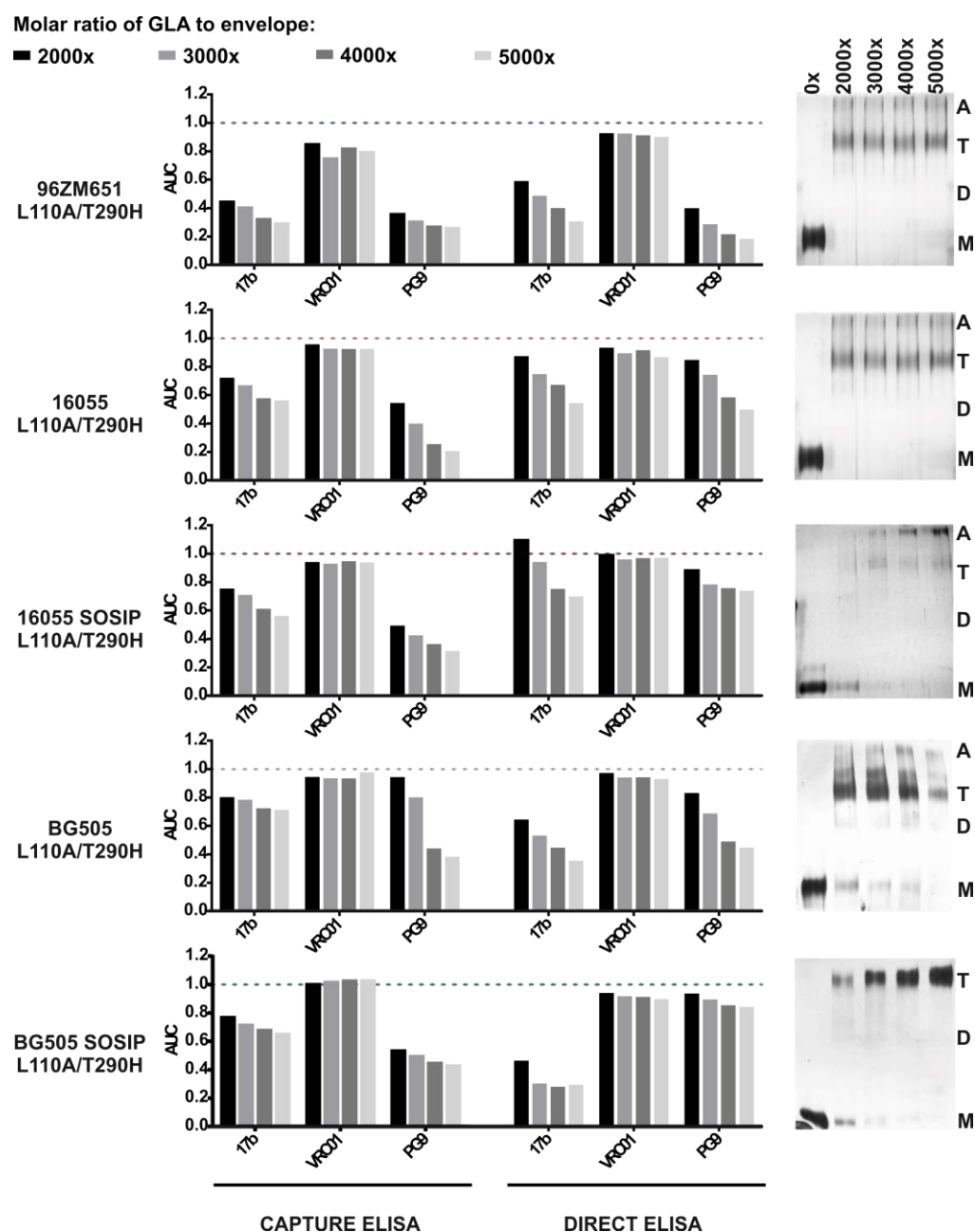
Figure 35: Antigenicity profiles of GLA cross-linked envelope trimers. GLA cross-linking was performed with 2000, 3000, 4000 and 5000 mol GLA per mol Env. Cross-linked trimers were analyzed by SDS-PAGE (right hand side; M: monomer, D: dimer, T: trimer, A: aggregate) and ELISA (left hand side) in a lectin capture format (preserves Env trimer structure) and direct coating format (mildly denaturing conditions promoting trimer opening and unfolding). Areas under the curve (AUC) were determined for VRC01 and PG9 binding to determine bnAb epitope integrity and maintenance of L111A/T278H gain of binding effects, as well as for CD4i antibody 17b, to analyze well-folded trimer integrity. AUCs were normalized to unmodified L111A/T278H variants (≙1). Although VRC01 signals could be maintained, PG9 binding was severely hindered at all concentrations where 17b recognition could be successfully diminished to ≤50 %. n=1.
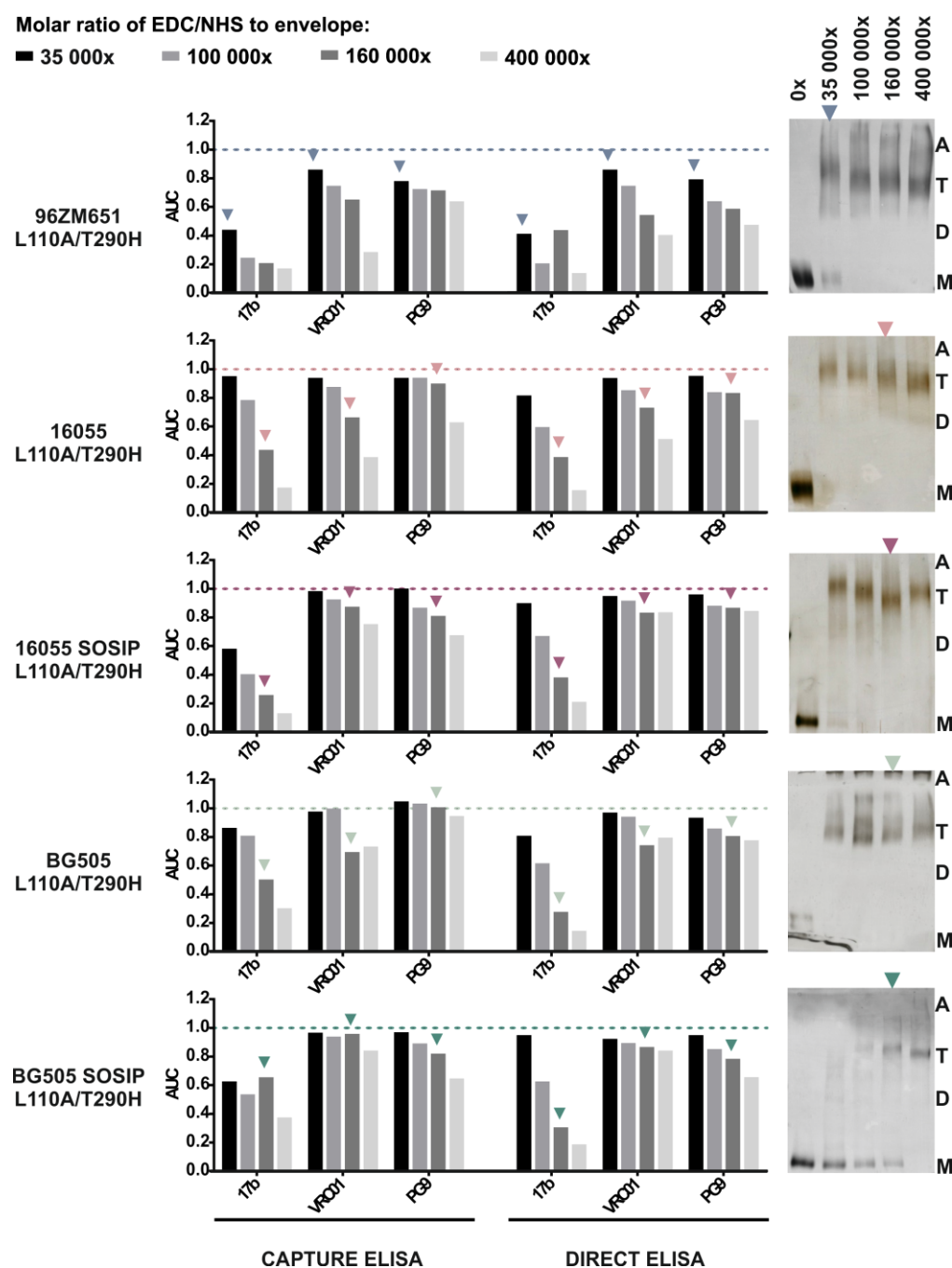
Figure 36: Antigenicity profiles of EDC/NHS cross-linked envelope trimers. EDC/NHS cross-linking was performed with 35 000, 100 000, 160 000 and 400 000 mol EDC/NHS per mol Env. Cross-linked trimers were analyzed by SDS-PAGE (right hand side; M: monomer, D: dimer, T: trimer, A: aggregate) and ELISA (left hand side) in a lectin capture format (preserves Env trimer structure) and direct coating format (mildly denaturing conditions promoting trimer opening and unfolding). Areas under the curve (AUC) were determined for VRC01 and PG9 binding to determine bnAb epitope integrity and maintenance of L111A/T278H gain of binding effects, as well as for CD4i antibody 17b, to analyze well-folded trimer integrity. AUCs were normalized to unmodified L111A/T278H variants ($\widehat{=}$1). Arrowheads indicate cross-linker concentrations where PG9 and VRC01 signals could be preserved ($\geq$75 %), while 17b binding was reduced ($\leq$50 %). n=1.

To further ensure that bnAb affinities of other epitopes on Env were sustained, L111A/T278H variants cross-linked with the respective optimal EDC/NHS concentration were analyzed with a broader antibody panel. This included the non-neutralizing or weakly neutralizing mAbs 17b (recognizing the open CD4i conformation) and F105 (targeting the CD4BS of aberrantly folded trimers), as well as bnAbs HGN194 (V3), PGT121 (V3-Glycan), HJ16 (CD4BS), VRC01 (CD4BS), PG9 (V1/V2), PGT145 (V1/V2) and PGT151 (gp120/gp41 interface). Using unmodified wildtype and L111A/T278H gp140 trimers as controls, cross-linked double mutants were analyzed by lectin capture and direct ELISA. Consistent with previous experiments (**Table 12**, **Figure 36**), L111A/T278H double mutants already displayed weaker affinity to 17b, which could be further reduced by EDC/NHS cross-linking. In addition, cross-linking severely reduced F105 binding for most isolates in both ELISA settings, further supporting the notion of a favorable effect of cross-linking on trimer integrity. As T278H removes the glycan at N276 necessary for HJ16 recognition, L111A/T278H double mutants showed impeded HJ16 binding. This, however, was not further influenced by cross-linking (**Figure 37A** and **Appendix Figure 6A**). Also in agreement with previous data (**Table 15**), insertion of L111A and T278H resulted in increased binding of VRC01 for most variants, which was not significantly influenced by cross-linking. Similarly, PG9 binding was slightly enhanced or preserved in the untreated double mutants and after cross-linking (**Figure 37B**). In addition, affinities to all other bnAb epitopes tested, including the V3 loop, were neither influenced by the two mutations, nor by subsequent cross-linking. This is consistent with previous results for L111A/T278H variants (**Table 12**) as well as literature findings regarding EDC/NHS cross-linking.[79] PGT145, a highly quaternary structure-dependent, apex-targeting bnAb,[58] and PGT151, which recognizes the gp120/gp41 interface and is therefore highly cleavage- and quaternary structure-dependent,[59] only showed noteworthy binding to BG505 SOSIP and 16055 SOSIP and not to the other cleavage-defective REKS mutants (**Figure 37B** and **C**).
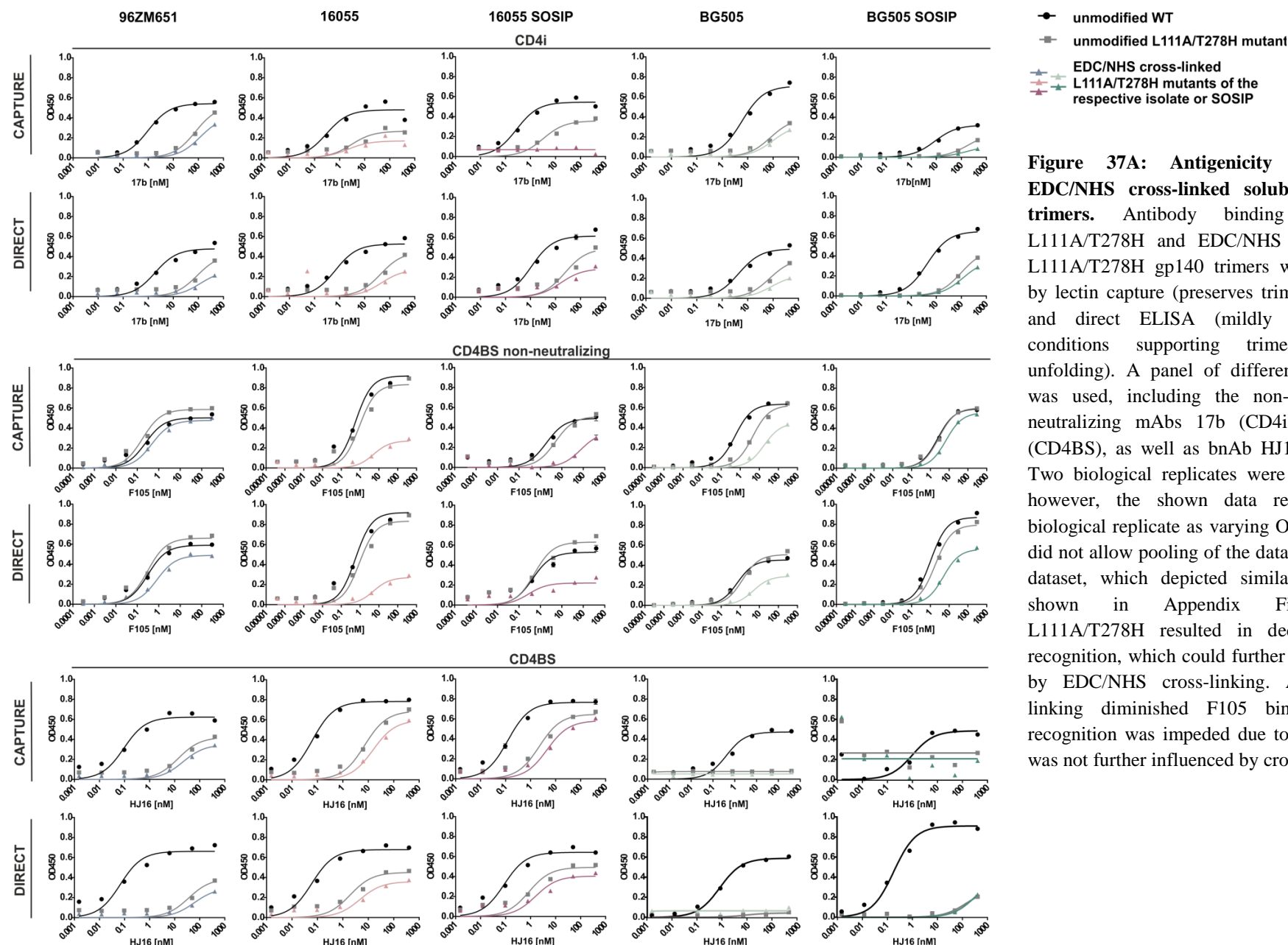
Figure 37A: Antigenicity profile of EDC/NHS cross-linked soluble envelope trimers. Antibody binding to WT, L111A/T278H and EDC/NHS cross-linked L111A/T278H gp140 trimers was analyzed by lectin capture (preserves trimer integrity) and direct ELISA (mildly denaturing conditions supporting trimer opening/ unfolding). A panel of different antibodies was used, including the non- or weakly neutralizing mAbs 17b (CD4i) and F105 (CD4BS), as well as bnAb HJ16 (CD4BS). Two biological replicates were carried out; however, the shown data represent one biological replicate as varying OD450 values did not allow pooling of the data. The second dataset, which depicted similar results, is shown in Appendix Figure 6A. L111A/T278H resulted in decreased 17b recognition, which could further be enhanced by EDC/NHS cross-linking. Also, cross-linking diminished F105 binding. HJ16 recognition was impeded due to T278H, but was not further influenced by cross-linking.
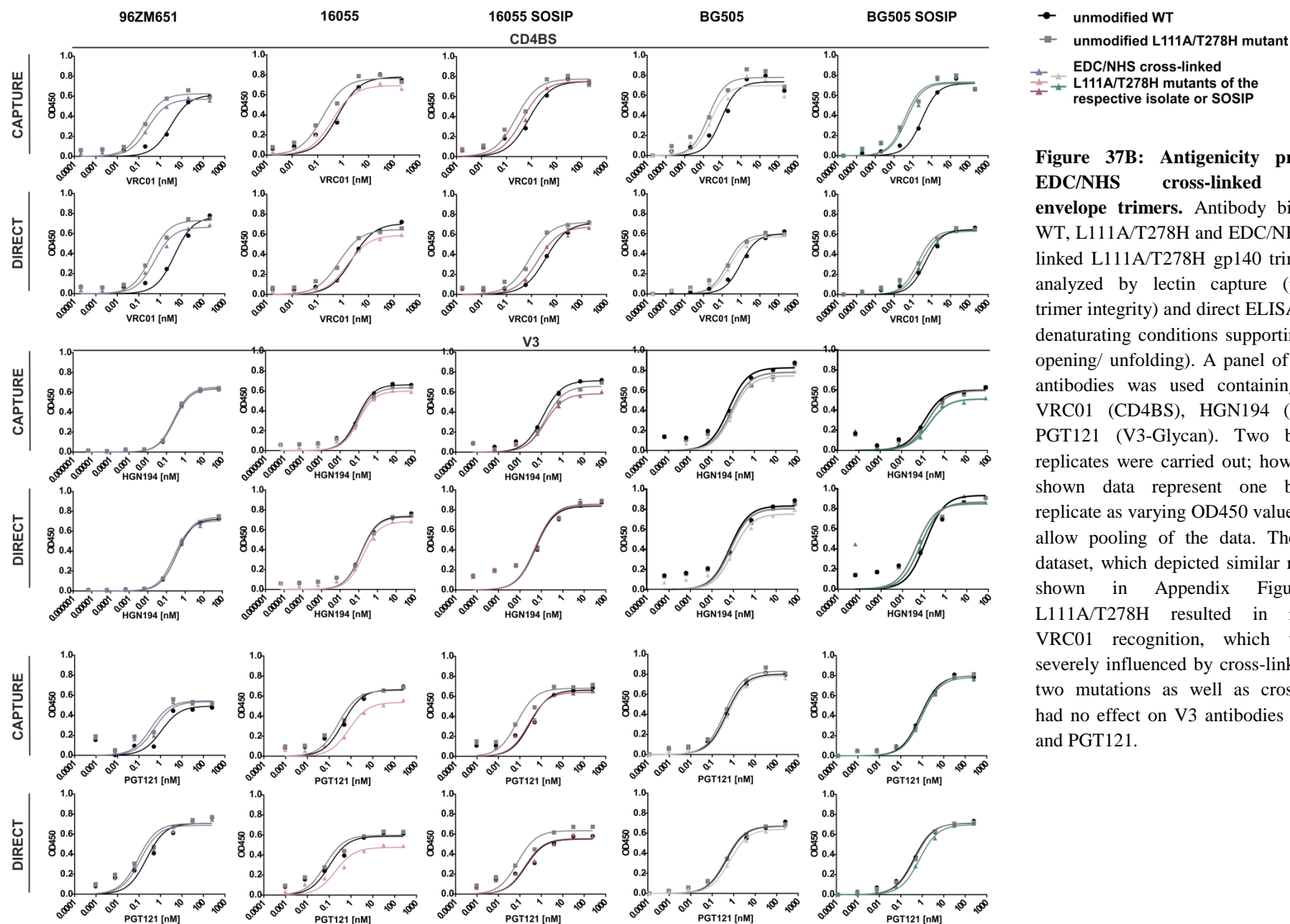
Figure 37B: Antigenicity profile of EDC/NHS cross-linked soluble envelope trimers. Antibody binding to WT, L111A/T278H and EDC/NHS cross-linked L111A/T278H gp140 trimers was analyzed by lectin capture (preserves trimer integrity) and direct ELISA (mildly denaturing conditions supporting trimer opening/ unfolding). A panel of different antibodies was used containing bnAbs VRC01 (CD4BS), HGN194 (V3) and PGT121 (V3-Glycan). Two biological replicates were carried out; however, the shown data represent one biological replicate as varying OD450 values did not allow pooling of the data. The second dataset, which depicted similar results, is shown in Appendix Figure 6B. L111A/T278H resulted in increased VRC01 recognition, which was not severely influenced by cross-linking. The two mutations as well as cross-linking had no effect on V3 antibodies HGN194 and PGT121.
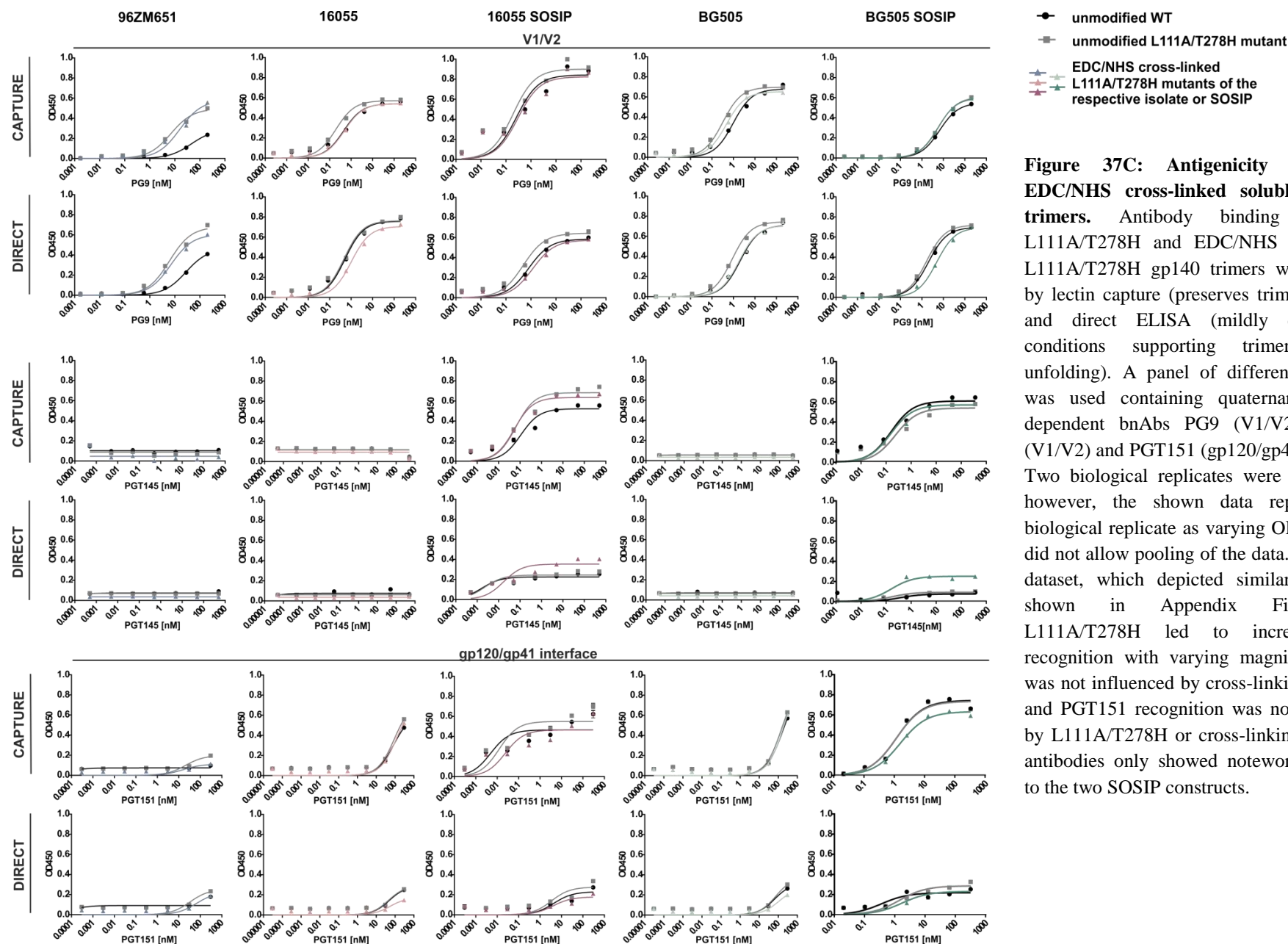
Figure 37C: Antigenicity profile of EDC/NHS cross-linked soluble envelope trimers. Antibody binding to WT, L111A/T278H and EDC/NHS cross-linked L111A/T278H gp140 trimers was analyzed by lectin capture (preserves trimer integrity) and direct ELISA (mildly denaturing conditions supporting trimer opening/ unfolding). A panel of different antibodies was used containing quaternary structure dependent bnAbs PG9 (V1/V2), PGT145 (V1/V2) and PGT151 (gp120/gp41 interface). Two biological replicates were carried out; however, the shown data represent one biological replicate as varying OD450 values did not allow pooling of the data. The second dataset, which depicted similar results, is shown in Appendix Figure 6C. L111A/T278H led to increased PG9 recognition with varying magnitude, which was not influenced by cross-linking. PGT145 and PGT151 recognition was not influenced by L111A/T278H or cross-linking, but these antibodies only showed noteworthy binding to the two SOSIP constructs.

Apart from the similarities of the different constructs regarding the effects of L111A/T278H and EDC/NHS cross-linking on antigenicity, all analyzed variants inherently depicted different affinity levels to the applied antibody panel (**Figure 38** and **Appendix Table 7**). 96ZM651, for example, showed markedly higher binding levels to F105, which recognizes aberrantly folded trimer,[133] although this was largely reduced by cross-linking. For BG505 and BG505 SOSIP, the unmutated wildtypes already exhibited a higher affinity to VRC01 than the other isolates, while 16055 and 16055 SOSIP showed higher recognition of PGT121 and marginally of PG9. PGT151 and PGT145, both highly dependent on proper trimer quaternary packing,[58,59] only recognized the well-folded engineered constructs BG505 SOSIP and 16055 SOSIP with satisfying affinity.
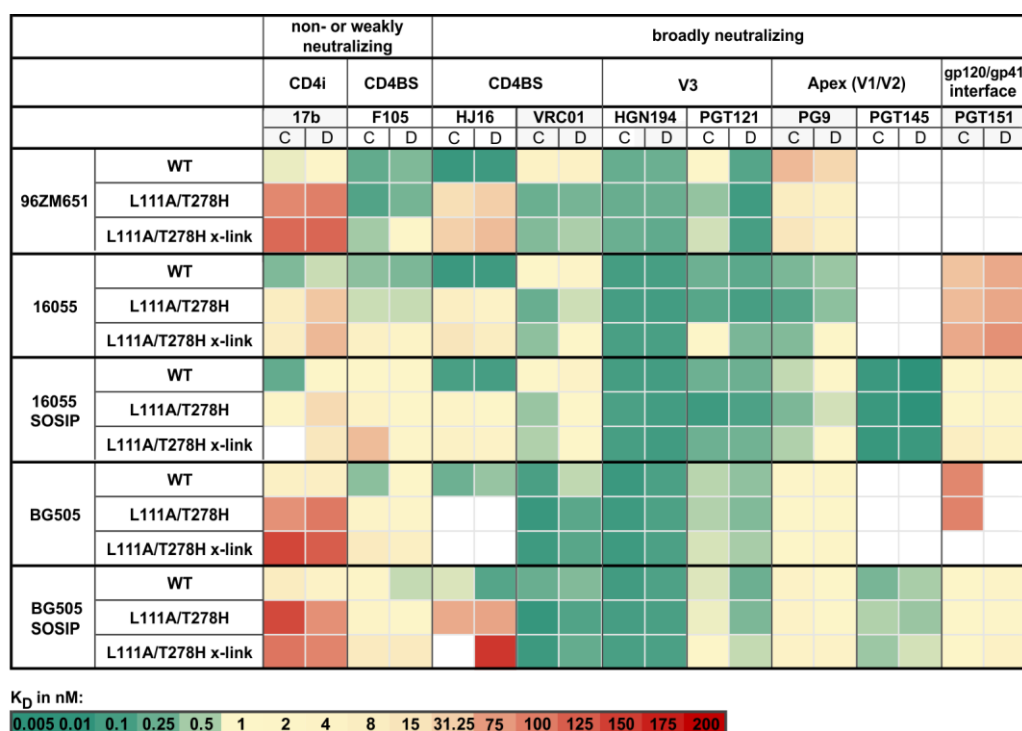
Figure 38: Antigenicity profiles of gp140 wildtypes, L111A/T278H mutants and EDC/NHS cross-linked L111A/T278H mutants of different isolates. K_D values (see Appendix Table 7) were calculated from titration curves of Figure 37 and Appendix Figure 6 (n=2). C: lectin capture ELISA, D: direct ELISA, CD4BS: CD4 binding site, CD4i: CD4-induced Env conformation. High to low K_D values are graded from red to green. White panels indicate non-calculable K_D values due to insufficient signal intensities (no binding).

In agreement with previous nanoDSF analyses of L111A single mutants (**Table 11**) and 96ZM651 combination mutants (**Appendix Table 6**), insertion of L111A/T278H did not significantly alter melting temperatures in any of the tested isolates. Cross-linking, however, increased thermostability for 16055 SOSIP and BG505 SOSIP, with cross-linked 16055 SOSIP exhibiting a T_M increase of approximately 6 °C (**Table 16**). Notably, melting curves of 96ZM651 allowed only inadequate analysis of melting temperatures due to low signal intensities, while 16055 could not be analyzed at all (**Appendix Figure 5**). This was most likely caused by inhomogeneous folding of the cleavage-incompetent REKS variants.

Table 16: Thermostability of cross-linked gp140 trimers. Thermostability of EDC/NHS cross-linked L111A/T278H trimers was analyzed by nanoDSF in comparison to unmodified WTs and L111A/T278H mutants (n=2). While insertion of L111A/T278H did not affect thermostability, cross-linking led to increased melting temperatures (T_M), most profoundly for 16055 SOSIP. Note that melting curves of 16055, BG505 and 96ZM651 were only poorly analyzable due to low signal intensities (Appendix Figure 5).

	T_M in °C		
	WT	**L111A/T278H**	**L111A/T278H x-link**
96ZM651	60.4 ± NA	61.1 ± 0.7	61.4 ± NA
16055	NA	NA	NA
16055 SOSIP	58.3 ± 0.3	59.6 ± 0.1	65.4 ± 0.1
BG505	58.1 ± 1.1	57.0 ± 0.0	61.1 ± 0.9
BG505 SOSIP	66.5 ± 0.1	67.0 ± 0.0	69.8 ± 0.2

In conclusion, 16055 SOSIP showed one of the more promising antigenicity profiles. In contrast to the cleavage-incompetent REKS variants, this included higher affinity to the quaternary structure-dependent bnAbs PG9 and PGT145, which target the trimer apex, as well as gp120/gp41 interface-targeting PGT151. In addition, 16055 SOSIP benefitted most from cross-linking regarding thermostability, which correlates with serum neutralization activity after immunization.[50]

4.2.8. Summary

In summary, application of the membrane-bound envelope alanine scanning library of clade C isolate 96ZM651 in a flow cytometry-based high-throughput assay enabled the identification of several mutations which improve the antigenicity profile of envelope immunogens. Thereby, W427A severely reduced CD4 recognition (**Table 6**), while mutants T278A/H and L111A exhibited increased signal intensities for bnAbs VRC01 and PG9, respectively (**Table 13**). Reduced CD4 binding would likely diminish the opening of the Env trimer, supporting a neutralizing antibody response *in vivo*. However, W427A also decreased the positive antigenicity effects of L111A and T278A/H (**Tables 12** and **14**), and was hence excluded from further analysis. Examination of the purified soluble 96ZM651 trimers revealed improved affinity of mutant T278A to VRC01. Interestingly, T278H and L111A gp140 trimers exhibited increased affinities to both VRC01 and quaternary-preferring PG9 (**Table 13**), hinting at potentially improved trimer folding for these variants. Additionally, L111A also enhanced envelope trimerization (**Figure 30**). Hence, T278H and L111A were analyzed in more detail in several membrane-bound and soluble variants of clades A, B and C, where both antigenicity and trimerization were also enhanced. Of note, this could also be demonstrated for isolates BG505 and 16055, as well as their trimer-stablilized SOSIP variants (**Tables 8** and **10**, **Figure 32**), which are currently extensively prevalent in the literature.[49,50,85,88]

Ultimately, the combination of T278A/H and L111A resulted in Env immunogens of different clades which show reduced recognition of non-neutralizing antibodies (**Table 12, Figure 37**), improved trimerization (**Figures 33** and **34**) as well as increased PG9 and VRC01 affinity. Specifically, soluble 96ZM651 L111A/T278H trimers exhibited six- to sevenfold enhanced affinity to PG9 and 47-fold improved affinity to VRC01, while the L111A/T278H double mutants of the other clades exhibited approximately twofold improved K_D values to PG9 and about fivefold enhanced affinities to VRC01 (**Tables 13** and **15**). Chemical cross-linking was able to further reduce the recognition of non-neutralizing antibodies (**Figures 37** and **38**) as well as increase envelope thermostability (**Table 16**).

Notably, the analyzed isolates inherently displayed varying magnitudes of affinities towards the applied antibody panel. Thereby, 96ZM651 exhibited markedly higher K_D values to both PG9 and VRC01 (**Tables 8, 10** and **15**), while also showing the weakest signals for quaternary-specific PGT145 and PGT151 (**Figure 38**). In contrast, 16055 and

16055 SOSIP L111A/T278H variants exhibited high affinities to the tested bnAb panel, including the quaternary-specific bnAbs (**Figure 38**). In addition, the 16055 SOSIP double mutant depicted the most profound improvement of trimerization (**Figure 34**) as well as the highest increase in thermostability after cross-linking (**Table 16**). Hence, 16055 L111A/T278H and 16055 SOSIP L111A/T278H, both with and without chemical cross-linking, will be further evaluated in a preclinical immunization study (**Figure 39**), to evaluate whether the herein demonstrated improved antigenicity and stability will ultimately translate into improved immunogenicity.

The most relevant effects of the mutations identified in this thesis for all tested constructs are summarized in **Table 17.**

Table 17: Summary of the effects of mutations L111A, T278A/H and W427A as well as their combinations on different isolates of clades A, B and C. The varians which are currently evaluated in a preclinical animal study are highlighted in red.

Clade	Isolate	Binding of gp145 and gp140 to quaternary-speific antibodies	Binding-FC[a] of mutant L111A to PG9	Binding-FC[a] of mutant L111A to VRC01	Binding-FC[a] of mutant T278X to VRC01 — T278A	Binding-FC[a] of mutant T278X to VRC01 — T278H	Preservation of PG9 and VRC01 GoB[a] effects in L111A/T278A/H double mutants	Binding-FC[a] of mutant W427A to soluble CD4	Binding-FC[a] of triple mutants L111A/T278A/H/ W427A to VRC01 and PG9	L111A and T278H were further evaluated in soluble trimers	Effect of L111A on trimerization[b]	Binding-FC[a] of mutant L111A to PG9 and VRC01 — PG9	Binding-FC[a] of mutant L111A to PG9 and VRC01 — VRC01	Binding-FC[a] of mutant T278H to VRC01	L111A/ T278H combination mutations — Binding-FC[a] to PG9 and (/) VRC01	L111A/ T278H combination mutations — Effect on trimerization (b)	Chemical cross-linking of L111A/ T278H double mutants — Preservation of PG9 and VRC01 GoB[a] effects	Chemical cross-linking of L111A/ T278H double mutants — Increase of T_M[a]
A	BG505	No	+	++	0	+	Yes	---			#	+	+	++	++ / ++++	0	Yes	NM
A	BG505 SOSIP	Yes	0	0	+	+++	Yes	---			*	++	+	+++	++ / +++	*	Yes	↑ (~ 3 °C)
B	HXB2	No	0	-	++	++	Yes	---	Reduced compared to double mutants		##	--	0	+++	NA	NA	Yes	NA
B	JRFL	No	-	0	++	++	Yes	---			##	+++	+	++	NA	NA	Yes	NA
C	16055	No	0	-	+	+	Yes	---			##	+	++	+++	+++ / ++++	##	Yes	NM
C	16055 SOSIP	Yes	+	0	+	++	Yes	--			###	0	0	++	+ / ++++	###	Yes	↑↑ (~ 6 °C)
C	96ZM651	No	++	+	+	+++	Yes	---			##	+	+	+++	++++ / ++++	##	Yes	NM

Fold changes for gp145 and gp140 mutants compared to WT binding levels:
++++: FC ≥ 5, +++: FC ≥ 2.5; ++: 2 ≤ FC < 2.5; +: 1.5 ≤ FC < 2; 0: 0.75 ≤ FC < 1.5; - : 0.5 ≤ FC < 0.75; -- : 0.1 ≤ FC < 0.5, --- : 0 ≤ FC <0.1

a) FC: Fold Change, GoB: Gain of Binding, T_M: Melting Temperature; NA: not available as measurements were not conducted; NM: not measurable due to too low signal intensity caused by aberrant trimer folding

b) Alterations in SEC elution profiles:
0: no effect; *: reduced expression yield; #: improved peak separation in SEC elution profiles ; ## : improved trimerization/ reduced monomer formation; ### : improved trimerization and increased trimer yield

5. Discussion

HIV`s vast diversity has impeded the success of common vaccine strategies like life attenuation or chemical inactivation.[16] Currently, HIV vaccine development focuses on the induction of broadly neutralizing antibodies, which target several different sites on the trimeric envelope protein.[33] However, Env`s immense diversity,[19,26] extensive glycan shield,[29] structural shielding of conserved sites[30,31] and structural instability[22] delay bnAb development up to several years during natural infection. Hence, induction of broadly neutralizing antibodies by active vaccination remains extremely difficult, requiring the development of improved envelope immunogens.

This thesis focused on the problem that both membrane-anchored and soluble Env trimers are flexible structures that exhibit different conformational states, influencing the presentation of structure-dependent bnAb epitopes and presenting non-neutralizing epitopes usually hidden in the pre-fusion closed state.[22,122] This is further aggravated when Env trimers bind to CD4 and undergo the CD4-induced conformational change.[49] Thus, Env variants with (i) improved bnAb binding, (ii) decreased recognition of non-neutralizing antibodies and CD4 as well as (iii) improved trimer structure and stability would likely support a cross-neutralizing humoral immune response. Here, an innovative screening technology for the identification of such Env variants according to their antibody binding properties was investigated. This system combined mammalian cell display of membrane-bound envelope variants with flow cytometry-based cell sorting to select envelopes with high or low affinity to the applied screening antibody. In addition, results obtained from analyzing envelope-bnAb and envelope-sCD4 interactions with an Env alanine scanning library were utilized for the rational design of novel Env immunogens with improved stability and antigenicity profiles.

5.1. Analysis of a mammalian cell display-based panning system for the high-throughput discovery of new AIDS vaccine candidates

5.1.1. Advantages of the applied stable cell line-based mammalian cell display system

In order to analyze the antigenic profiles of HIV-1 envelope variants in the described panning process, recombinant Env proteins are displayed on inducible, stably transfected HEK cell lines, which are widely used e.g. for producing pseudoviruses for neutralization assays.[55,54,134] This provides several advantages over other currently used methods. In contrast to yeast display of HIV-1 envelopes,[135,136] it not only allows the expression of functional Env trimers in their natural membrane context, but also ensures mammalian glycosylation and native folding. Second, contrary to screening approaches with soluble trimers,[137,138,139] this approach also enables the assessment of untruncated envelope variants, which was recently shown to have an important influence on Env folding and antigenicity.[65] Hence, this display technique should provide all proposed requirements of envelope presentation, including glycan-dependent epitopes on Env,[55,45,58] the preservation of promising quaternary structure-dependent epitopes[59,58] as well as membrane-proximal epitopes like the MPER domain,[140,48] which is e. g. deleted in soluble trimer-stabilized SOSIP variants.[49]

Notably, the genomic integration of library vector pQL13 at a single, distinct FRT site allows a highly stringent linkage of genotype and phenotype within the library. In addition, the defined integration locus also reduces expression variations that could potentially be caused by the genomic context (such as promotors or regulatory regions). Residual deviations in expression levels of the different variants can be efficiently normalized by the translational coupling of Env and eGFP by a "self-cleaving" TaV 2A peptide[104] (**Figure 13**).

In addition, the eGFP/Env expression cassette is under inducible regulation of the Tat operator/repressor system, preventing Env cytotoxicity effects during cell cultivation and thus, variations in library composition caused by growth advantages or disadvantages of

specific variants. Moreover, if multiple panning cycles have to be applied to achieve satisfying enrichment rates, this stable cell line-based screening approach also allows the simple re-expansion of a portion of the sorted cells, while the remaining cells can be subjected to further enrichment rate analysis.

5.1.2. Single-round FACS panning of the Env/V3 model library

Taken together, the above described stringent genotype-phenotype linkage and expression normalization accounted for high enrichment rates for both low and high affinity variants in only a single round of panning, as exemplified using a five-member chimeric model library and mAbs 447-52D and HGN194. Notably, high affinity chimera Env/V3-MN could be selectively enriched to represent 71 - 93 % of all sequences out of an equimolar mixture of the five chimeras, resulting in enrichment rates of up to 56- and 55-fold for 447-52D and HGN194, respectively. Similarly, the abundance of low affinity variant Env/V3-SF33 for 447-52D and the two non-binders Env/V3-SF33 and -RF for HGN194 was significantly increased in the respective low affinity fractions: Env/V3-SF33 showed an up to 237-fold enrichment in the low affinity gate for 447-52D (72 - 97 % out of an equimolar mixture of all five chimeras), while the two low affinity variants for HGN194, Env/V3-SF33 and -RF, were enriched approximately 5-fold (ca. 50 % each out of an equimolar mixture of all five chimeras) (**Table 1** and **Figure 16**). Importantly, we were able to replicate those results using qPCR, capillary sequencing of single clones as well as next generation sequencing on both the re-cloned plasmid DNA as well as directly on the genomic DNA of sorted cells and input controls (**Figure 16**), although enrichment rates varied slightly between the different methods (**Table 1**). However, the utilized formula[108] not only considers the enrichment of the respective selected variants, but also the depletion of all other chimeras and is thus highly sensitive to slight variations between the different analyses.

In conclusion, this procedure represents a strong improvement over the panning approach previously published by Bruun *et al.*, which utilized a lentiviral vector system for stable transgene integration. There, even a low multiplicity of infection (MOI) of 0.1 could not prevent multiple infections and thus multiple integration events in single cells. The resulting lack of stringent genotype-phenotype linkage made two panning cycles necessary

to obtain 81 % Env/V3-MN sequences out of an equimolar mixture of all five variants, representing an enrichment of only 20-fold (estimated by qPCR).[96]

5.1.3. Future directions

Notably, next generation sequencing was successfully applied directly to the genomic DNA of sorted cells and input controls (**Figure 16**), enabling the facilitated and efficient analysis of large libraries. Hence, this approach would be easily applicable to more comprehensive trimeric Env libraries and highly potent broadly neutralizing antibodies like VRC01,[141] PG9/PG16[142] or members of the PGT family.[55,59] Including available engineered unmutated common bnAb ancestors (UCAs) into the screening, this system could also be extended to identify Env variants for sequential immunization approaches.[143,144,145,146] Such a strategy was e.g. recently reported with envelope variants identified from a panning approach which used lentiviral transgene integration and subsequent FACS sorting,[147] similar to the approach published by Bruun *et al.*[96] The problem of multiply infected cells was thereby apparently solved by extended cultivation of stable cell lines and MOIs of less than 0.1. Using different BG505 SOSIP gp140 based libraries fused to a PDGFR-transmembrane domain, Steichen *et al*. engineered several HIV-1 Env trimers with affinity for germline PGT121-class bnAbs, intermediates with increasing levels of epitope modification and mature envelope variants. Sequential immunization with these soluble trimers effectively induced bnAbs in two PGT121 knock-in mouse models,[84] demonstrating the practical applicability of such a panning technology. In contrast to this approach, the here described strategy offers the possibility to include membrane-proximal bnAb epitopes by administering the isolated membrane-bound Env immunogens via non-viral or viral vectors.

In addition to the membrane-bound variants, antigenicity profiles of the model library to 447-52D and HGN194 were also investigated for soluble gp140, with comparable results (**Figure 17**, **Table 2**). Although the magnitude of K_D values varied between flow cytometry of gp145 variants and ELISA analysis of gp140 trimers, most probably due to different inherent conditions of the two methods, the overall order of binding affinities for each investigated epitope was equal. Hence, vaccine candidates selected with this approach may, similarly to the BG505 SOSIP-based constructs described above,[147,84] also be potent

immunogens as soluble Env trimers. To this end, however, secreted variants would have to be evaluated individually, as phenotypes might not always be easily transferable from membrane-bound to soluble trimers, e.g. where epitopes on gp41 are concerned.

5.1.4. Conclusion

In conclusion, we have provided an efficient mammalian cell display and panning system that allows the selection of complete trimeric HIV envelopes. The stable cell line-based display of membrane-bound Env variants provides the advantages of (i) inducible Env expression, (ii) mammalian folding and glycosylation, (iii) efficient normalization for Env expression via GFP fluorescence and (iv) stringent genotype-phenotype linkage due to single integration into the predefined FRT site (**Figure 13**). Data generated for the five-member model library demonstrate a clear correlation between (i) the MFIs measured at a distinct 447-52D or HGN194 antibody concentration for the individual membrane-bound Env/V3 chimeras (**Figure 14A** and **C**), (ii) gp145 binding affinities to 447-52D or HGN194 (**Table 2**), (iii) the copy number (qPCR) or number of sequences (capillary sequencing / NGS) of the respective Env/V3 chimeras after one panning cycle (**Figure 16**) and (iv) binding affinities of 447-52D or HGN194 to soluble Env/V3 trimers (**Table 2**). The limits regarding feasible library sizes still have to be explored in greater detail, which is currently further investigated by Julia Koop as part of her PhD thesis. However, this technology may ultimately be applicable to diverse viral vaccines, including not only HIV, but also other bnAb-dependent vaccine endeavors like Influenza,[148,19] Hepatitis C virus[149,150] or Cytomegalovirus, where the proper assembly of several envelope subunits is necessary for the induction of highly quaternary structure-dependent broadly neutralizing antibodies.[151] Although not specifically addressed in this thesis, this panning system may additionally be useful for the rapid epitope mapping of yet uncharacterized antibodies. To this end, alanine scanning libraries of envelope surface residues or glycan scanning libraries with modified N-glycosylation sites or patterns could be employed to investigate the footprints of mature bnAbs and their engineered germline variants.

5.2. Development of HIV-1 envelope immunogens with improved antigenicity and stability using a cell-based epitope mapping platform

5.2.1. Epitope mapping with the 96ZM651 alanine scanning library

Recent progress in sorting patient B cell populations with engineered well-folded envelope baits[152,52,153] and high-throughput neutralization scans of immortalized B cell repertoires[54,48,44] has led to the identification of various new broadly neutralizing antibodies. This has resulted in an increased demand for methods which facilitate and accelerate the fine mapping of antibody-envelope interfaces to gather knowledge for the rational design of new Env immunogens. While peptide scans only yield information about linear epitopes[44] and are therefore not suitable for most of the newly identified structure-dependent bnAbs,[42,58,60] Env-antibody co-crystallization reveals the interaction at atomic level resolution.[42,110] However, some of the available bnAb structures were not solved by Env co-crystallization, but in an unbound state, like for example PGT145 (PDB ID: 3UIS)[117] or PGDM1400 (PDB ID: 4RQQ).[51] In addition, some antibodies proved to be difficult to crystallize, e.g. HJ16, which was only recently co-crystallized with Env (PDB ID: 4YE4, Chen *et al.*, to be published). In this thesis, a mammalian cell display and flow cytometry-based high-throughput alanine scanning technique was employed to map envelope interactions with sCD4, CD4 binding site-targeting bnAb VRC01[42] and the quaternary structure preferring bnAb PG9[58] targeting the trimer apex.

5.2.1.1. General setup and evaluation of the alanine scanning procedure

Display of Env alanine variants on HEK cells ensured native folding and mammalian glycosylation, providing all proposed requirements of Env presentation (see 5.1.1) to facilitate the analysis of glycan[45,55,58] or structure-dependent[58,59,60] as well as MPER-targeting[48,140] bnAbs.

Since the well-folded SOSIP and NFL variants, which are better suited for the analysis of conformation-dependent antibodies, have only recently become available,[49,82] a cleavage-incompetent gp145 envelope of clade C isolate 96ZM651 was used to avoid gp120 shedding.[116] It is now known that cleavage site mutation disturbs Env quaternary structure;[81] however, sCD4 and VRC01 both recognize soluble gp120[20,42] and PG9 is quaternary-preferring,[127] but was reported to also bind several gp120 proteins and cleavage-incompetent gp140 monomers and trimers.[49,57] Here, these investigated interaction partners were able to tolerate cleavage site mutation in comparison to the cleavage-competent wildtype (**Figure 19**). However, other newly identified bnAbs like PGT145,[58] PGT151-158[59] or 35O22[60] are heavily reliant on well-folded envelope trimers and did not bind cleavage-defective 96ZM651 (**Figure 37**), prohibiting mapping approaches of highly structure-dependent mAbs with this library. To overcome these limitations, fully cleaved alanine scanning libraries based on BG505 SOSIP[49,154] gp145 (resulting in enhanced Env surface expression levels[64]), and gp160 (excluding antigenicity effects of cytoplasmic tail deletion[64,65]) are currently being established in the course of Benjamin Zimmer`s PhD thesis within the group.

To normalize for varying Env surface presentation levels and transfection efficiencies, co-staining with mAb 5F3 was implemented, which recognizes a linear epitope in gp41.[46] This limits the mapping technology mostly to the analysis of gp120-targeting screening reagents, due to sterical constraints caused by simultaneous staining with two antibodies. To overcome these limitations, a panel of different reference antibodies (including e.g. HGN194, which recognizes a linear epitope in V3,[44] or 2G12, which targets a conserved mannose-patch in gp120[45]) could be utilized to include all alanine variants into the assay and enable the characterization of gp41-targeting bnAbs. Moreover, normalizing reagents of smaller sizes would also diminish sterical hindrances. To this end, Fab fragments could be applied, as well as heavy chain antibodies of shark or camelid origin[155,156] or Env-targeting aptamers[157,158], although such reagents would be more complex to obtain. Moreover, Steichen *et al.*[147] inserted a c-Myc tag into different BG505 SOSIP gp140 based libraries, located between gp140 and a PDGFR transmembrane domain. As this strategy was already proven to effectively normalize for Env expression in FACS sorting experiments, it could also be employed in future libraries, e.g. by inserting a linker-tag-linker sequence between amino acid position 664, where gp140 SOSIP constructs are terminated,[49] and the HIV-1 envelope transmembrane domain.

5.2.1.2. Analysis of envelope interactions with sCD4, VRC01 and PG9

All alanine scans conducted in this thesis were able to map the interaction between 96ZM651 Env and the analyzed ligands (sCD4, VRC01 and PG9) to the respective targeted region. However, the screenings did not result in a 100 % overlap with published crystal structure data (**Figures 23, 25** and **28**), possibly due to an interplay of several reasons. First, while co-crystallization detects the whole interaction surface of two binding partners, not all contact residues actually contribute to the interaction energy. Mutational studies therefore confirmed that epitopes defined structurally differ from epitopes identified in functional assays.[124] On the other hand, alanine substitutions do not induce extreme electrostatic or steric effects and hence, do not significantly alter a protein`s main chain conformation. Therefore, alanine scanning libraries allow the identification of amino acid side chain interactions, but main-chain-only contacts are likely to be unaffected.[111,125]

This, for example, can explain the negative influence on sCD4 binding for mutations N425A and W427A, where main chain and side chain contacts account for the interaction. In contrast, neighbouring residues like M426, Q428 and G431 were not influenced, as only main chain interactions play a role at these positions (**Figure 23A**).[42]

In addition to contacts determined by crystallography, all alanine scans identified further significant loss of binding mutations, often located on the inside of the envelope trimer and highly conserved (**Appendix Table 3**). Hence, these mutations supposedly have a negative influence on Env structure, resulting in decreased binding for all tested interaction partners, although PG9 targets a completely different epitope compared to sCD4 and VRC01. This is in agreement with previous reports stating that e.g. mutations F92A or M99A inside the envelope trimer disrupt Env conformation.[159] The generation of more extensive datasets, including different envelope epitopes and highly structure-dependent mAbs like PGT151[59] or PGT145[58] could help to exclude such positions from future epitope mapping analyses.

5.2.2. Development of next generation envelope immunogens

5.2.2.1. Identification of significant mutations for the rational design of next generation envelope immunogens

Interestingly, the performed alanine scans not only allowed characterization of the respective epitopes, but also yielded mutations which can be directly introduced into next generation envelope immunogens. T278A and H were found to enhance VRC01 binding in the flow cytometry-based assay, with fold changes of approximately twofold and four- to fivefold, respectively (**Figure 26**, **Table 7**). In addition, L111A led to about twofold increased signal intensities for PG9 (**Tables 9** and **10**). Notably, these effects of T278A/H and L111A were also transferable to soluble gp140 trimers (**Tables 8** and **10**).

Both T278 mutants result in the disruption of an N-glycosylation site at position 276 at the edge of the CD4 binding site (**Figure 25**). This glycan was previously shown to be essential for bnAb HJ16[160] (**Figure 37**). However, its removal also results in increased neutralization of other CD4BS-targeting antibodies, for example bnAb NIH45-46 and germline reverted versions of NIH45-46 or VRC01,[144] supposedly by improving accessibility of the CD4 binding site. Apart from glycan removal, the additional benefit of T278H might be explained by advantageous interactions of histidine`s positive charge with VRC01 (**Figure 26B**). However, T278H not only positively affected VRC01 binding with fold changes of about fivefold for membrane-bound Env and an approximately threefold improved K_D for soluble gp140, but also enhanced binding of quaternary-preferring bnAb PG9 to both gp145 and gp140 (**Table 13**). As PG9 targets the trimer apex and can therefore not directly interact with position 278 at the CD4 binding site, this might suggest improved folding of mutant T278H, supporting both tertiary structure representation of the CD4 binding site epitope as well as quaternary packing of the trimer.

L111A was previously described to decrease aberrant gp120 dimerization as well as recognition of sCD4 and CD4i antibodies, thereby improving the neutralizing antibody response.[129,130] Recently, BG505 L111A gp120 has been implemented in several immunization studies in comparison to trimeric BG505 SOSIP.[49,80,84] This thesis demonstrated that L111A not only affects gp120, but also impacts envelope trimerization, as demonstrated by the profoundly reduced monomer fraction of gp140 L111A (**Figure**

32). Furthermore, L111A positively influenced trimer antigenicity. While the previously described effect of impaired sCD4 and CD4i binding could be replicated for 96ZM651 membrane-bound gp145 and soluble gp140, both also depicted increased binding to structure-dependent bnAb VRC01 and quaternary structure-dependent PG9 (**Tables 9** and **13**). However, position L111 is located in the inner domain of Env, at the distal end of the α1 helix[161] and is therefore not available for direct antibody interactions. Interestingly, though, this particular region was shown to play a role in the conformational change leading to the CD4 bound state of gp120.[161] In conjunction with the more homogeneous melting curve observed by DSC analysis of L111A gp140 (**Figure 31**), this hints at a favorable impact of L111A on envelope structure. Thereby, Env quaternary folding may be improved. However, as PG9 also binds distinct gp120 and gp140 monomers[49,57] and reduced binding of CD4 and CD4i mAbs was also observed for gp120,[129,130] L111A could also result in improved folding of the gp120 subunits, while trimer integrity might not be affected. To properly elucidate the proposed structural influence of both T278H and L111A on HIV Env, negative-stain electron microscopy[81] could be applied to the respective soluble trimers in comparison to wildtype gp140.

Notably, the effects of L111A and T278A/H as originally determined for 96ZM651 envelope (clade C) were also well transferable to other gp145 and gp140 variants of clades A, B and C, including trimer-stabilized SOSIP constructs (**Figures 32** and **34, Tables 8** and **10**). Thereby, T278H exhibited a slight advantage over T278A in all tested gp145 variants, while increasing VRC01 affinity to the soluble trimers approximately two- to threefold (**Table 8**). Most likely due to the different effects of L111A (i.e. on PG9 affinity, Env trimerization and possibly Env structure), fold changes in signal intensities for membrane-bound L111A variants and K_D values for secreted trimers were not conform for all isolates. However, most tested soluble L111A variants displayed approximately twofold increased affinity to both PG9 and VRC01 (**Table 10**), as well as markedly improved trimerization. For the latter, the only exception presented BG505 and BG505 SOSIP. This may be explained by the inherent predisposition of BG505 to form stable trimers,[162,81] which is further enhanced by the SOSIP mutations[49] and might not be additionally improvable by L111A. Notably, the most pronounced effect regarding envelope trimerization was observed for 16055 SOSIP, where introduction of L111A resulted in a markedly increased trimer fraction (**Figure 32**).

In conclusion, the identified gain of binding mutations improved affinities of bnAbs VRC01 and PG9 as well as envelope trimerization. In addition, both T278H and L111A

may additionally affect trimer folding, which could further be elucidated with negative-stain electron microscopy.[81] Originally identified for membrane-bound 96ZM651, the beneficial phenotypes were not only transferable to soluble 96ZM651 trimers, but also to membrane-bound and soluble variants of different isolates and SOSIP constructs. This reinforces the notion that results obtained from the 96ZM651 based CSAM procedure are also of value for other applications and can provide indications for the development of soluble vaccine candidates from different subtypes. Antigenicity profiles were, however, not invariably applicable to all tested constructs. For instance, L111A did reduce the monomer fraction of soluble gp140 in B clade isolate HXB2, but nevertheless impaired PG9 affinity (**Figure 32**, **Table 10**). Thus, the potential of mutations which were identified in 96ZM651 gp145 on other membrane-bound or soluble envelope isolates would always have to be evaluated individually.

5.2.2.2. Combination of antigenicity improving single mutations

Membrane-bound gp145 and soluble gp140 Env proteins containing both L111A and T278A/H mutations resulted in next generation envelope immunogens with increased bnAb binding, decreased recognition of non-neutralizing antibodies and improved envelope trimerization (**Tables 12** and **13, Figures 33** and **34**). Most interestingly, the effects of L111A and T278H thereby seemed to be additive, resulting in six- to sevenfold improved affinity of double mutant L111A/T278H to PG9 and 47-fold improved affinity to VRC01 in the case of 96ZM651. In addition, the K_D values of L111A/T278H variants based on other Env isolates were reduced approximately twofold for PG9 and fivefold for VRC01 (**Table 13**, **Table 15**). In agreement with results observed for the L111A single mutants, double mutants of BG505 and the associated SOSIP construct again exhibited no effect regarding trimerization, while the most prominent impact was observed for 16055 SOSIP, where the trimer fraction was markedly increased in SEC elution profiles (**Figure 34**).

Furthermore, the alanine scanning procedure revealed significantly impaired binding of W427A to sCD4 and the non-neutralizing antibody 17b, which targets the open, CD4-induced envelope conformation (**Appendix Table 4**). The highly conserved tryptophan residue is located in the bridging sheet of Env[163] and directly interacts with both VRC01

and CD4 (**Figures 23** and **25**).[42] However, while VRC01 contacts only its main chain, CD4 forms both main chain and side chain interactions.[42] Hence, we were able to optionally include W427A into the next generation Env immunogens, resulting in preserved bnAb binding, while recognition of sCD4 and 17b was severely inhibited (**Table 12**). Abolishing CD4 binding would supposedly support antigen bioavailability and hence a neutralizing antibody response, as envelope immunogens could not be depleted by binding to CD4$^+$ host immune cells before they had the opportunity to interact with B-cell receptors or be taken up by antigen-presenting cells.[79] More importantly, however, reduced CD4 recognition may be relevant to prohibit the CD4-induced opening of the Env trimer and the resulting presentation of potentially distractive non-neutralizing epitopes *in vivo*.[164] However, to thoroughly analyze these effects on immunogenicity, a human CD4 transgenic[165,166,167] or humanized animal model[168] would have to be used. As W427A in both gp145 and gp140 also slightly decreased VRC01 and PG9 binding compared to L111/T278 double mutants in both gp145 and gp140 (**Tables 12** and **13**), it was, for the time being, excluded from further analysis.

5.2.2.3. Chemical cross-linking of L111A/T278H envelope variants

Since combination of L111A and T278H resulted in the most prominent increase in VRC01 and PG9 recognition (**Tables 13** and **14**), we decided to investigate the potential for further stabilization of L111A/T278H variants by chemical cross-linking. To this end, double mutants of 96ZM651, BG505/BG505 SOSIP and 16055/16055 SOSIP were used. Cross-linking of vaccine antigens was first applied to create inactivated polio vaccines and has since been used more widely with an excellent safety record.[169,170] Here, we utilized a recently published protocol for HIV-1 envelope cross-linking with glutaraldehyde (GLA), which cross-links amine groups by inserting a five-member carbon-chain. In addition, EDC/NHS was applied, which cross-links carboxyl- and amine-groups without inserting additional atoms.[78,79,50] For GLA-treated antigens, extensive safety data are available for clinical trials of heart[171] and vascular[172] transplants, as well as allergy immunotherapy.[173,174] However, GLA cross-linking severely impaired PG9 binding for all variants (**Figure 35**) and was thus excluded from further analysis. This is in line with a

recent report of Schiffner *et al.*, who found that GLA cross-linking drastically decreased binding of PGT145, which shares the core epitope of PG9.[58]

In contrast to GLA, no clinical trials have been conducted with EDC/NHS cross-linked proteins to date, however, as it does not insert additional linker atoms into the immunogen, safety issues seem unlikely.[175] Supposedly due to different cross-linking chemistries and the absence of linker atoms, EDC/NHS could, in contrast to GLA, successfully stabilize envelope trimers. While presentation of non-neutralizing epitopes for all L111A/T278H variants was further reduced, bnAb binding to various epitopes could largely be preserved (**Figure 37**). Notably, contrary to Schiffner *et al.*,[79] who hypothesized that cross-linked trimers can only display one of the two quaternary structure-dependent epitopes for PGT145 and PGT151 efficiently, the epitopes of both antibodies could be maintained.

Due to the inhomogeneous folding of cleavage-incompetent gp140 trimers and the resulting low signal intensities for melting curves of 96ZM651, BG505 and 16055, thermostability of cross-linked trimers could only reliably be determined for the two SOSIP constructs. In agreement with recent reports,[79,50] both cross-linked SOSIP constructs depicted increased melting temperatures (**Table 16**), which was proven to correlate with improved immunogenicity.[50]

In summary, the reduced presentation of non-neutralizing epitopes and increased stability of cross-linked variants may contribute to improved neutralizing antibody responses *in vivo*, which will further be analyzed in a preclinical immunization study (**Figure 39**).

5.2.3. Conclusion

In conclusion, the described alanine scanning procedure is a fast and effective method to characterize antibody-Env interactions. It provides the advantages of native folding and mammalian glycosylation and can, in combination with other techniques like co-crystallization, supply additional information about bnAb footprints. Furthermore, this approach is able to directly yield mutations for the rational design of envelope immunogens, as demonstrated with L111A and T278A/H, which both increased VRC01 and PG9 affinity. While L111A was proven to additionally enhance envelope trimerization, both L111A and T278H might also improve trimer folding. Originally identified for membrane-bound 96ZM651, these mutations also enhanced antigenicity and

supposedly the structure of soluble 96ZM651 trimers as well as membrane-bound and secreted variants from different clades. Hence, this screening approach is able to guide the development of novel membrane-bound and soluble vaccine candidates from different subtypes.

Combination of the identified mutations resulted in envelope immunogens with reduced recognition of non-neutralizing antibodies, improved trimerization as well as increased bnAb affinity. Specifically, mutations L111A and T278H appeared to have an additive effect, leading to six- to sevenfold improved affinity of PG9 and 47-fold improved affinity of VRC01 for 96ZM651. Of note, L111A/T278H double mutants from other clades exhibited approximately twofold reduced K_D values for PG9 and about fivefold decreased K_Ds for VRC01. These favourable antigenicity profiles could additionally be enhanced by chemical cross-linking, which further reduced recognition of non-neutralizing antibodies and increased thermostability, which should both support a neutralizing antibody response *in vivo*.[50]

Notably, the different isolates inherently displayed different magnitudes of affinities towards the applied bnAb panel. Thereby, 96ZM651 exhibited markedly higher K_D values to both PG9 and VRC01 compared to BG505/BG505 SOSIP and 16055/16055 SOSIP (**Tables 8, 10** and **15**), while also showing the weakest signals for quaternary-specific PGT145 and PGT151 (**Figure 38**). Interestingly, the trimer-stabilized BG505 SOSIP repeatedly exhibited lower affinity to both VRC01 and PG9 than the unmodified clade A isolate (**Tables 8, 10** and **15**), which was similarly reported for CD4BS bnAbs VRC01 and b12.[81] In summary, 16055 and 16055 SOSIP L111A/T278H variants exhibited one of the more promising antigenic profiles, including binding of quaternary-specific bnAbs (**Figure 38**). In addition, the 16055 SOSIP double mutant depicted the highest increase in thermostability after cross-linking as well as the most profound improvement of trimerization. While thermostability was recently shown to correlate with serum neutralization capacity after vaccination,[50] a shift towards a higher trimer proportion on the surface of Env producing cells will likely also support induction of a more potent, quaternary-specific antibody response after DNA priming. Hence, these variants were chosen to be further evaluated in a preclinical immunization study in rabbits, to analyze whether the demonstrated improved antigenicity and stability will ultimately translate into enhanced immunogenicity.

6.   Perspective

The herein described flow cytometry-based mammalian cell display and panning approach was evaluated with a five-member chimeric model library. As a next step, the procedure has to be expanded to larger and more sophisticated libraries. For this, a sequential permutation library of 16055 gp145 is available. Analyses of the panning technique with five, 15 and 45 permutated positions, i.e. a maximum of 100, 300 or 900 different variants, was part of Julia Koop`s PhD thesis. These sublibraries were exclusively composed of variants with mutations at the CD4BS of Env. Of note, each library contained the permutated position 278, where T278A and H were shown to increase VRC01 binding in this thesis. Preliminary results suggest that both mutants could be isolated in the high affinity gate of the five positions screen after one panning cycle, while the superior T278H was also enriched from the sublibrary containing 15 positions. Other variants from the high and low affinity gate are currently tested by flow cytometry to verify their increased or decreased binding to VRC01. Eventually, the screening should be applicable to the complete 16055 sequential permutation library of approximately 13,000 variants. Furthermore, this panning approach is intended to be the basis of projects within the EHVA (European HIV Vaccine Association) consortium. In particular, an envelope library based on the well-folded BG505 SOSIP construct is to be established, to identify germline-targeting envelope variants for the highly effective bnAb PGT145 for sequential immunization approaches. Other projects in cooperation with Dr. Rogier Sanders at the Amsterdam Medisch Centrum (AMC), who first published BG505 SOSIP,[49] include the increase of expression yields or improved trimer folding for not yet published SOSIP versions. To be able to screen these extensive envelope libraries, the library generation procedure as well as the sorting strategy will have to be adapted to be applicable to a large number of different variants.

The 16055 and 16055 SOSIP variants identified with the alanine scanning technique will be applied in a preclinical immunization study to survey whether improved antigenicity and stability *in vitro* will translate into improved immunogenicity *in vivo*. Formulation of all antigens will occur in GLA-LSQ (Glucopyranosyl Lipid Adjuvant-Liposome QS21), which was already effective for vaccine candidates of leishmania, influenza or malaria.[176,177,178] Recently, a formulation study in cooperation with the Infectious Diseases

Research Institute (IDRI), Seattle, USA, demonstrated physicochemical stability of all envelope antigens for the intended rabbit study with this adjuvant.

Each immunization arm will consist of five rabbits. The first four groups will receive protein primes and boosts of one protein each (either 16055 WT, 16055 L111A/T278H, 16055 SOSIP WT or 16055 SOSIP L111A/T278H), to evaluate the influence of the L111A/T278H and SOSIP mutations on immunogenicity. Three further groups will receive DNA primes and protein boosts of 16055 and 16055 SOSIP variants in different combinations. DNA priming and the resulting expression of envelope variants on the surface of affected host cells will be valuable to elucidate the impact of the profoundly improved trimerization of the 16055/16055 SOSIP double mutants. In all these groups, protein stability will increase with the time point of administration to analyze whether stabilized immunogens are needed as priming or boosting reagents to support a neutralizing antibody response. For all seven groups, the fourth and last boost will consist of the respective cross-linked envelope trimer to determine whether this further stabilization can additionally improve neutralizing activity or maintenance of the humoral immune response. An overview of the immunization study, which is anticipated to start in June 2017, is shown in **Figure 39**.
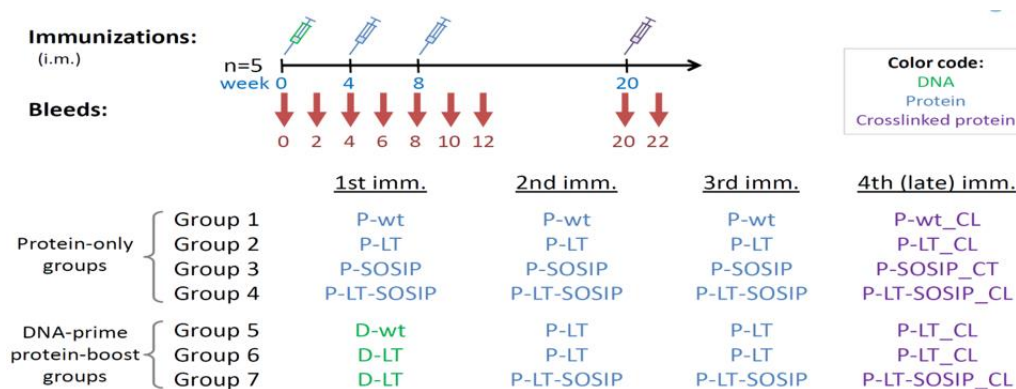
Figure 39: Schematic overview of the preclinical immunization study. Upper panel: New Zealand White rabbits will receive four immunizations at weeks 0, 4, 8 and 20. Red arrows indicate samplings of sera for the analysis of antibody titers (ELISA) and neutralization capacity against tier 1 and tier 2 viruses (TZM-bl or A3R5 assays). Lower panel: Groups 1-4 will receive protein (P) primes and boosts of one gp140 trimer each (16055 WT, 16055 L111A/T278H (LT), 16055 SOSIP WT or 16055 SOSIP L111A/T278H). Groups 5-7 will receive DNA (D) primes and protein boosts of different combinations of 16055 and 16055 SOSIP variants, with trimer stability increasing with the time point of administration. The last boost of each group will contain the respective cross-linked (CL) envelope trimer.

7. Abstract

HIV`s vast genetic diversity has prevented the success of common vaccine approaches. Currently, HIV-1 vaccine development focuses on the induction of broadly neutralizing antibodies (bnAbs), which target various epitopes on the trimeric envelope (Env) protein. The most potent bnAbs known today are able to neutralize more than 90 % of circulating virus strains; however, their development requires a complex co-evolution of virus and humoral immune system. Therefore, induction of bnAbs by active immunization remains extremely difficult, demanding both improved Env immunogens as well as innovative screening techniques for their identification.

The first project of this thesis investigated a mammalian cell display and cell sorting-based panning technique for the identification of vaccine candidates with improved antigenicity. This panning procedure combines the advantages of (i) single integration of Env into a distinct FRT site to link genotype and phenotype within a stable cell line library, (ii) inducible Env expression to avoid cytotoxicity effects, (iii) translational coupling of Env and GFP to normalize for induced Env expression and (iv) display on HEK cells to ensure native folding and mammalian glycosylation. Using a model library of five envelope chimeras with distinct binding profiles to mAbs 447-52D and HGN194, cell sorting selectively enriched the high affinity variant up to 55- and 56-fold, respectively, and low affinity variants up to 237-fold after only a single round of panning. Notably, affinity rankings of cell surface displayed and purified secreted trimers were equal, suggesting that binding properties of Env variants selected with this approach can also guide the identification of soluble envelope antigens. Applying more elaborate Env libraries to screen the currently available set of broadly neutralizing monoclonal antibodies and their engineered unmutated common ancestors, this stable cell line-based panning technology therefore provides a valuable tool for the selection of new HIV-1 Env vaccine candidates. Ultimately, this technique may also be applicable to other bnAb-dependent vaccine endeavors targeting e.g. Influenza-, Hepatitis C- or Cytomegalovirus.

A second project focused on the mapping of envelope-bnAb interactions for the rational design of novel Env immunogens. To this end, mammalian cell display of an envelope alanine library based on C clade isolate 96ZM651 was applied in a high-throughput flow cytometry assay to map envelope interactions with different mAbs as well as soluble CD4 (sCD4). Interestingly, this approach was able to directly identify amino acid substitutions

which improved Env antigenicity and folding. Combination of these mutations ultimately resulted in both membrane-bound and soluble envelope immunogens with increased affinity to bnAbs PG9 (targeting the trimer apex) and VRC01 (targeting the CD4 binding site), while the binding of non-neutralizing mAb 17b could be profoundly reduced. Notably, these trimers also exhibited enhanced trimerization and potentially improved folding. Optionally, a mutation which drastically reduced recognition of CD4 could be included, which could be beneficial for a neutralizing antibody response *in vivo* by prohibiting the profound CD4-induced conformational change of Env. Although the improved structural and antigenic phenotypes of those mutants were originally identified for 96ZM651 envelope, these favourable properties were also applicable to variants of other clade A, B and C isolates, including engineered, trimer-stabilized SOSIP constructs. Of note, membrane-bound gp145 and soluble gp140 variants thereby exhibited the same effects, indicating that this flow cytometry-based mapping procedure is able to guide the design of membrane-bound and soluble vaccine candidates from different subtypes. Additionally, affinity-enhanced soluble envelope trimers were stabilized by chemical cross-linking. This further reduced the binding of non-neutralizing antibodies and markedly increased thermostability, which correlates with the elicitation of neutralizing antibodies. The most potent envelope immunogens developed in this thesis will be analyzed in a preclinical immunization study starting in June 2017 to determine whether their improved antigenicity and stability *in vitro* will ultimately translate into enhanced immunogenicity *in vivo*.

8. Zusammenfassung

Aufgrund der enorm hohen genetischen Variabilität von HIV-1 sind konventionelle Impfstoffstrategien bislang gescheitert. Heute fokussiert sich die HIV-1 Vakzineentwicklung vor allem auf die Induktion breitneutralisierender Antikörper (bnAbs), die sich gegen verschiedene Epitope auf dem trimeren Oberflächenprotein Envelope (Env) richten. Die Besten heute bekannten bnAbs können über 90 % der zirkulierenden Virusstämme neutralisieren. Ihre Entstehung erfordert allerdings eine komplexe Koevolution zwischen Virus und humoralem Immunsystem, sodass ihre Induktion durch aktive Immunisierung immer noch enorm schwierig ist. Aus diesem Grund sind sowohl verbesserte Env Immunogene, als auch Methoden zu deren Identifizierung nötig.

Das erste Projekt dieser Promotion beschäftigte sich mit einer Panning-Technologie, mit der durch Zellsortierung Envelopes mit verbesserten Antigenitätsprofilen aus einer Bibliothek basierend auf stabilen Zelllinien selektiert werden können. Dabei werden folgende Vorteile miteinander kombiniert: (i) Integration nur einer Env Variante in eine definierte FRT-Stelle pro Zelle, um Genotyp und Phänotyp innerhalb der Bibliothek effektiv zu verlinken, (ii) induzierbare Env Expression, um Zytotoxizitätseffekten vorzubeugen, (iii) translationale Verknüpfung von GFP und Env, um auf die induzierte Env Expression normalisieren zu können und (iv) Expression der Env Varianten auf HEK-Zellen, um native Faltung und Säugetierglykosylierung zu gewährleisten. Unter Verwendung einer Modelbibliothek aus fünf Env-Chimären mit definierten Bindungsprofile zu den monoklonalen Antikörpern 447-52D und HGN194 konnte die Variante mit der höchsten Affinität zu beiden Antikörpern in nur einem Panningzyklus jeweils bis zu 55- bzw. 56-fach angereichert werden. Gleichzeitig zeigten die jeweiligen niedrig affinen Varianten Anreicherungsfaktoren von bis zu 237-fach. Interessanterweise zeigten membranständige und gereinigte lösliche Env Trimere die gleiche Rangfolge bezüglich ihrer Affinitäten. Somit kann diese Technologie auch für die Identifizierung löslicher Antigene wegweisend sein. Mit Hilfe komplexerer Envelope-Bibliotheken und dem derzeit vorhandenen Set an bnAbs und deren berechneter Keimbahnvarianten stellt diese Methode somit ein wertvolles Instrument zur Identifizierung neuer HIV Vakzinekandidaten dar. Dabei ist dieses Verfahren nicht nur auf HIV beschränkt, sondern

kann auch auf andere bnAb-abhängige Impfstoffvorhaben ausgeweitet werden, wie z.B. Influenza, Hepatitis C oder Cytomegalievirus.

Das zweite Projekt dieser Promotion beruhte auf der Kartierung von Envelope-bnAb-Interaktionen für das rationale Design neuer Env Immunogene. Dazu wurde eine membranständige Envelope Alanin-Bibliothek des Subtyp C Isolats 96ZM651 in einem Durchflusszytometrie-basierten Hochdurchsatzverfahren verwendet. Die Expression der Varianten fand dabei ebenfalls auf HEK-Zellen statt, um native Faltung und Säugetierglykosylierung zu gewährleisten. Mit Hilfe dieser Technologie wurden die Interaktionsflächen von Env sowohl zu den breitneutralisierenden Antikörpern PG9, welcher den Apex des Trimers angreift, und VRC01, der die CD4 Bindestelle erkennt, als auch zu löslichem CD4 (sCD4) untersucht. Interessanterweise wurden dabei Mutationen identifiziert, welche sowohl die Antigenität, als auch die Trimerisierung und Faltung von Env positiv beeinflussen. Die Kombination dieser Mutationen resultierte schließlich in erhöhter Affinität zu PG9 und VRC01, verringerter Bindung des nicht neutralisierenden Antikörpers 17b und verbesserter Trimerbildung. Optional konnte durch Einfügen einer weiteren Mutation außerdem die Bindung an sCD4 verhindert werden. Da diese eine tiefgreifende Konformationsänderung in Env auslöst, welche eine Vielzahl nicht neutralisierender Antikörperepitope freilegt, würde dies *in vivo* wahrscheinlich eine neutralisierende Antikörperantwort unterstützen. Die ursprünglich für 96ZM651 Envelope identifizierten Mutationen und deren Kombinationen konnten erfolgreich auf andere Isolate der Subtypen A, B und C, sowie auf trimerstabilisierte SOSIP-Varianten übertragen werden. Außerdem zeigten membranständige und lösliche Envelope Trimere die gleichen Effekte, was den Wert dieses Verfahrens für die Entwicklung löslicher Vakzinekandidaten unterschiedlicher Subtypen verdeutlicht.

Außerdem konnten die generierten Mutanten durch chemisches Cross-Linking weiter stabilisiert werden. Dies reduzierte zusätzlich die Bindung nicht-neutralisierender Antikörper, während die Thermostabilität der Trimere, und somit deren Fähigkeit neutralisierende Antikörperantworten zu induzieren, deutlich erhöht wurde. Eine präklinische Immunisierungsstudie wird ab Juni 2017 prüfen, ob die *in vitro* demonstrierte verbesserte Antigenität und Stabilität der generierten Immunogene letztendlich auch deren Immunogenität positiv beeinflusst.

9. Appendix

Appendix Table 1: Oligonucleotide sequences used in this thesis.

Primer Name	Primer Sequence	Primer Name	Primer Sequence
Mutagenesis primer for the insertion of point mutations into different HIV-1 isolates			
96ZM651 L110A for	GGACATCATCAGCGCCTGGGACCAGAGCC	**zm96 T290Q rev**	GGTTTTCACGTTGTTCTGCAGGTTCTCGCTTC
96ZM651 L110A rev	GGCTCTGGTCCCAGGCGCTGATGATGTCC	**zm 96 T290E for**	GAAGCGAGAACCTGGAGAACAACGTGAAAACC
96ZM651 S162A for	CGACATGAAGAACTGCGCCTTCAACATCACCAC	**zm 96 T290E rev**	GGTTTTCACGTTGTTCTCCAGGTTCTCGCTTC
96ZM651 S162A rev	GTGGTGATGTTGAAGGCGCAGTTCTTCATGTCG	**zm96 T290K for**	GAAGCGAGAACCTGAAGAACAACGTGAAAAC
96ZM651 S162D for	CGACATGAAGAACTGCGACTTCAACATCACCACC	**zm96 T290K rev**	GTTTTCACGTTGTTCTTCAGGTTCTCGCTTC
96ZM651 S162D rev	GGTGGTGATGTTGAAGTCGCAGTTCTTCATGTCG	**96ZM651 L110A W111A for**	GGACATCATCAGCGCCGCCGACCAGAGCC
96ZM651 S162T for	CATGAAGAACTGCACCTTCAACATCACCAC	**96ZM651 L110A W111A rev**	GGCTCTGGTCGGCGGCGCTGATGATGTCC
96ZM651 S162T rev	GTGGTGATGTTGAAGGTGCAGTTCTTCATG	**MG_REKS>2xG4S_Fwd**	ATATGGTCTCACTGGCGGAGGGGGAAGTGCCGTGGGCATCGGCGCCGTG
96ZM651 F163D for	CATGAAGAACTGCAGCGACAACATCACCACCGAG	**MG_REKS>2xG4S_Rev**	ATATGGTCTCACCAGATCCGCCTCCGCCCTCCACCACCCGCCGCTTGGCC
96ZM651 F163D rev	CTCGGTGGTGATGTTGTCGCTGCAGTTCTTCATG	**BG505 L110A for**	CCGACATCATCAGCGCCTGGGACCAGAGCC
96ZM651 N164M for	GAACTGCAGCTTCATGATCACCACCGAGCTG	**BG505 L110A rev**	GGCTCTGGTCCCAGGCGCTGATGATGTCGG
96ZM651 N164M rev	CAGCTCGGTGGTGATCATGAAGCTGCAGTTC	**BG5O5 L110W for**	CCGACATCATCAGCTGGTGGGACCAGAGCC
96ZM651 N164S for	GAACTGCAGCTTCAGCATCACCACCGAG	**BG5O5 L110W rev**	GGCTCTGGTCCCACCAGCTGATGATGTCGG
96ZM651 N164S rev	CTCGGTGGTGATGCTGAAGCTGCAGTTC	**BG505 V126A for**	GACCCCCCTGTGCGCCACCCTGCAGTGCAC
96ZM651 N164V for	GAAGAACTGCAGCTTCGTGATCACCACCGAGCTG	**BG505 V126A rev**	GTGCACTGCAGGGTGGCGCACAGGGGGGTC
96ZM651 N164V rev	CAGCTCGGTGGTGATCACGAAGCTGCAGTTCTTC	**BG505 L128A for**	CCTGTGCGTGACCGCCCAGTGCACCAACG
96ZM651 E168I for	CAGCTTCAACATCACCACCATCCTGAAGGACAAGAAAAAG	**BG505 L128A rev**	CGTTGGTGCACTGGGCGGTCACGCACAGG
96ZM651 E168I rev	CTTTTTCTTGTCCTTCAGGATGGTGGTGATGTTGAAGCTG	**BG505 N160A for**	GGGCGAGCTGAAGGCCTGCAGCTTCAAC
96ZM651 E168Q for	CAACATCACCACCCAGCTGAAGGACAAGAG	**BG505 N160A rev**	GTTGAAGCTGCAGGCCTTCAGCTCGCCC
96ZM651 E168Q rev	CTTGTCCTTCAGCTGGGTGGTGATGTTG	**BG505 S162A for**	GAGCTGAAGAACTGCGCCTTCAACATGACCAC
96ZM651 E168T for	CAGCTTCAACATCACCACCACCCTGAAGGACAAGAAAAAG	**BG505 S162A rev**	GTGGTCATGTTGAAGGCGCAGTTCTTCAGCTC
96ZM651 E168T rev	CTTTTTCTTGTCCTTCAGGGTGGTGGTGATGTTGAAGCTG	**BG5O5 S162T for**	CGAGCTGAAGAACTGCACCTTCAACATGACCACCG
96ZM651 K172C for	CACCGAGCTGAAGGACTGCAAAAAGAACGTGTACGC	**BG5O5 S162T rev**	CGGTGGTCATGTTGAAGGTGCAGTTCTTCAGCTCG
96ZM651 K172C rev	GCGTACACGTTCTTTTTGCAGTCCTTCAGCTCGGTG	**BG505 F163A for**	CTGAAGAACTGCAGCGCCAACATGACCACCG
96ZM651 K172L for	CACCGAGCTGAAGGACCTGAAAAAGAACGTGTAC	**BG505 F163A rev**	CGGTGGTCATGTTGGCGCTGCAGTTCTTCAG
96ZM651 K172L rev	GTACACGTTCTTTTTCAGGTCCTTCAGCTCGGTG	**BG505 N164A for**	GAACTGCAGCTTCGCCATGACCACCGAGC
96ZM651 K172P for	CACCGAGCTGAAGGACCCCAAAAAGAACGTGTACG	**BG505 N164A rev**	GCTCGGTGGTCATGGCGAAGCTGCAGTTC
96ZM651 K172P rev	CGTACACGTTCTTTTTGGGGTCCTTCAGCTCGGTG	**BG505 T166A for**	CAGCTTCAACATGGCCACCGAGCTGC
96ZM651 K173I for	GAGCTGAAGGACAAGATCAAGAACGTGTACGCC	**BG505 T166A rev**	GCAGCTCGGTGGCCATGTTGAAGCTG
96ZM651 K173I rev	GGCGTACACGTTCTTGATCTTGTCCTTCAGCTC	**BG505 E168A for**	CTTCAACATGACCACCGCCCTGCGGGACAAGAAAC
96ZM651 K173L for	CGAGCTGAAGGACAAGCTGAAGAACGTGTACGCC	**BG505 E168A rev**	GTTTCTTGTCCCGCAGGGCGGTGGTCATGTTGAAG
96ZM651 K173L rev	GGCGTACACGTTCTTCAGCTTGTCCTTCAGCTCG	**BG505 K172A for**	CCGAGCTGCGGGACGCCAAACAGAAGGTGTAC
96ZM651 T290A for	CAGAAGCGAGAACCTGGCCAACAACGTGAAAAC	**BG505 K172A rev**	GTACACCTTCTGTTTGGCGTCCCGCAGCTCGG

Primer Name	Primer Sequence	Primer Name	Primer Sequence
96ZM651 T290A rev	GTTTTCACGTTGTTGGCCAGGTTCTCGCTTCTG	BG505 K173A for	GAGCTGCGGGACAAGGCCCAGAAGGTGTACAG
zm96 T290R for	GAAGCGAGAACCTGAGGAACAACGTGAAAAC	BG505 K173A rev	CTGTACACCTTCTGGGCCTTGTCCCGCAGCTC
zm96 T290R rev	GTTTTCACGTTGTTCCTCAGGTTCTCGCTTC	BG505 I185A for	CTACCGGCTGGACGCCGTGCAGATCAACG
zm96 T290Q for	GAAGCGAGAACCTGCAGAACAACGTGAAAACC	BG505 I185A rev	CGTTGATCTGCACGGCGTCCAGCCGGTAG
BG505 T290A for	CAGAAGCGAGAACATCGCCAACAACGCCAAGAAC	16O55 T290A rev	GTTTTCACGTTGTTGGCCAGGTTCTCGCTTCTG
BG505 T290A rev	GTTCTTGGCGTTGTTGGCGATGTTCTCGCTTCTG	16O55 T290H for	GAAGCGAGAACCTGCACAACAACGTGAAAAC
BG5O5 T290H for	CAGAAGCGAGAACATCCACAACAACGCCAAGAAC	16O55 T290H rev	GTTTTCACGTTGTTGTGCAGGTTCTCGCTTC
BG5O5 T290H rev	GTTCTTGGCGTTGTTGTGGATGTTCTCGCTTCTG	16O55 N426A for	GATCAAGCAGATCATCGCCATGTGGCAAGAAGTG
BG5O5 N426A for	GATCAAGCAGATCATCGCCATGTGGCAGCGGATCGG	16O55 N426A rev	CACTTCTTGCCACATGGCGATGATCTGCTTGATC
BG5O5 N426A rev	CCGATCCGCTGCCACATGGCGATGATCTGCTTGATC	16O55 W428A for	CAGATCATCAATATGGCCCAAGAAGTGGGCAGAG
BG5O5 W428A for	GCAGATCATCAATATGGCCCAGCGGATCGGCCAGGC	16O55 W428A rev	CTCTGCCCACTTCTTGGGCCATATTGATGATCTG
BG5O5 W428A rev	GCCTGGCCGATCCGCTGGGCCATATTGATGATCTGC	16055 SOSIP W428A for	GATCATTAACATGGCCCAGGAAGTGGGCAG
BG505 REKR to REKS fw	GTGGGACGGGAGAAGAGGGCCGTGGGCATCGG	16055 SOSIP W428A rev	CTGCCCACTTCCTGGGCCATGTTAATGATC
BG505 REKR to REKS rev	CCGATGCCCACGGCCCTCTTCTCCCGTCCCAC	16055 SOSIP T290H for	CATCAGAAGCGAGAATCTGCACAACAATGTGAAAACCATC
16055 L110A for	GGACGTGATCAGCGCCTGGGACCAGAGCC	16055 SOSIP T290H rev	GATGGTTTTCACATTGTTGTGCAGATTCTCGCTTCTGATG
16055 L110A rev	GGCTCTGGTCCCAGGCGCTGATCACGTCC	HXB2 L110A for	GGACATCATCAGCGCCTGGGACCAGAGCC
16O55 L110W for	GACGTGATCAGCTGGTGGGACCAGAG	HXB2 L110A rev	GGCTCTGGTCCCAGGCGCTGATGATGTCC
16O55 L110W rev	CTCTGGTCCCACCAGCTGATCACGTC	HXB2 L110W for	GACATCATCAGCTGGTGGGACCAGAG
16055 V126A for	GACCCCCCTGTGCGCCACCCTGGAATGCAG	HXB2 L110W rev	CTCTGGTCCCACCAGCTGATGATGTC
16055 V126A rev	CTGCATTCCAGGGTGGCGCACAGGGGGGTC	HXB2 S162T for	GATCAAGAACTGCACCTTCAACATCAGC
16055 L128A for	CCCTGTGCGTGACCGCCGAATGCAGACAGG	HXB2 S162T rev	GCTGATGTTGAAGGTGCAGTTCTTGATC
16055 L128A rev	CCTGTCTGCATTCGGCGGTCACGCACAGGG	HXB2 T290A for	GCGTGAACTTCGCCGACAACGCC
16055 N160A for	GCGAGGAAATCAAGGCCTGCAGCTTCAATGC	HXB2 T290A rev	GGCGTTGTCGGCGAAGTTCACGC
16055 N160A rev	GCATTGAAGCTGCAGGCCTTGATTTCCTCGC	HXB2 T290H for	GAAGCGTGAACTTCCACGACAACGCCAAG
16055 S162A for	GAAATCAAGAACTGCGCCTTCAATGCCACCACC	HXB2 T290H rev	CTTGGCGTTGTCGTGGAAGTTCACGCTTC
16055 S162A rev	GGTGGTGGCATTGAAGGCGCAGTTCTTGATTTC	HXB2 N426A for	GATCAAGCAGATCATCGCCATGTGGCAGAAAGTG
16O55 S162T for	TCAAGAACTGCACCTTCAATGCCACC	HXB2 N426A rev	CACTTTCTGCCACATGGCGATGATCTGCTTGATC
16O55 S162T rev	GGTGGCATTGAAGGTGCAGTTCTTGA	HXB2 W428A for	CAGATCATCAATATGGCCCAGAAAGTGGGCAAG
16055 F163A for	CAAGAACTGCAGCGCCAATGCCACCACCG	HXB2 W428A rev	CTTGCCCACTTTCTGGGCCATATTGATGATCTG
16055 F163A rev	CGGTGGTGGCATTGGCGCTGCAGTTCTTG	JRFL L110A for	GAAGATATCATCAGCGCCTGGGACCAGAGCCTG
16055 N164A for	CAAGAACTGCAGCTTCGCCGCCACCACCGAGATC	JRFL L110A rev	CAGGCTCTGGTCCCAGGCGCTGATGATATCTTC
16055 N164A rev	GATCTCGGTGGTGGCGGCGAAGCTGCAGTTCTTG	JRFL L110W for	GATATCATCAGCTGGTGGGACCAGAGC
16055 T166A for	CAGCTTCAATGCCGCCACCGAGATCC	JRFL L110W rev	GCTCTGGTCCCACCAGCTGATGATATC
16055 T166A rev	GGATCTCGGTGGCGGCATTGAAGCTG	JRFL S162T for	GATCAAGAACTGCACCTTCAACATCACC
16055 E168A for	CTTCAATGCCACCACCGCCATCCGGGACAAGAAAC	JRFL S162T rev	GGTGATGTTGAAGGTGCAGTTCTTGATC
16055 E168A rev	GTTTCTTGTCCCGGATGGCGGTGGTGGCATTGAAG	JRFL T290A for	GCGACAACTTCGCCAACAACGCC
16055 K172A for	CACCGAGATCCGGGACGCCAAACAGAAGGTGTAC	JRFL T290A rev	GGCGTTGTTGGCGAAGTTGTCGC
16055 K172A rev	GTACACCTTCTGTTTGGCGTCCCGGATCTCGGTG	JRFL T290H for	GAAGCGACAACTTCCACAACAACGCCAAG

Primer Name	Primer Sequence	Primer Name	Primer Sequence
16055 K173A for	GAGATCCGGGACAAGGCCCAGAAGGTGTACGC	JRFL T290H rev	CTTGGCGTTGTTGTGGAAGTTGTCGCTTC
16055 K173A rev	GCGTACACCTTCTGGGCCTTGTCCCGGATCTC	JRFL N426A for	CAAGCAGATCATTGCCATGTGGCAGGAAAG
16055 I185A for	GTTCTACCGGCTGGACGCCGTGCCCCTGGAAGAG	JRFL N426A rev	CTTCCTGCCACATGGCAATGATCTGCTTG
16055 I185A rev	CTCTTCCAGGGGCACGGCGTCCAGCCGGTAGAAC	JRFL W428A for	CAGATCATTAACATGGCCCAGGAAGTGGGCAAG
16O55 T290A for	CAGAAGCGAGAACCTGGCCAACAACGTGAAAAC	JRFL W428A rev	CTTGCCCACTTCCTGGGCCATGTTAATGATCTG
Primer Name	PrimerSequence	Primer Name	PrimerSequence
Cloning primer for the insertion of envelope variants into pWLXLd and different QL vectors			
pWPXLd 96ZM651 for BamHI	ACGGGATCCCGTCTCGCTAGCCGG	pWPXLD 16055 rev EcoRI	CGGGAATTCGTCTCCTCGAGCCTGCAGGTCATCAGCTGTATCCCTGCCGCACTCTGTTCACCAC
pWPXLd 96ZM651 rev EcoRI	CGGGAATTCGTCTCCTCGAGCCTGCAG	pWPXLD 16O55 for	GAGGTTTAAACTACGGGATCCCGTCTCG
pWPXLd zm96 for	ACGGGATCCCGTCTCGCTAGCCGGACCGGCCACCATGGGAGTGCGGGAGATCCTGCG	pWPXLD 16O55 rev	CGGGAATTCGTCTCCTCGAGCC
pWPXLd zm96 rev	CGGGAATTCGTCTCCTCGAGCCTGCAGGTCATCAGGCGTAGTCGGGCACATCG	16055-QL-For	ATAATAGCTAGCCGTCTCCCTAGCATGAGAGTGCGGGGCATCCTGCG
96ZM651 QL for	ATAATACGTCTCGCTAGCATGGGAGTGCGGGAGATCCTGCGGAACTGGCAGCGGTGGTGG	16055-gp145 QL Rev	ATATATACTCGAGCGTCTCGTCGAGTCATCAGCTGTATCCCTGCCGC
96ZM651 gp145 QL rev	ATATTCGTCTCCTCGAGCTAGTAGCCCTGCCGCACTCTGTTCACGATGGACAGCACGGCG	16055-gp145 HA-Tag QL Rev	ATATATCGTCTCCTCGAGTTATCATCAGGCGTAGTCGGGCACGTCGTAGGGGTAGTATCCCTGCCGCACTCTGTTCACC
96ZM651 QL for2	ATAATACGTCTCGCTAGCATGGGAGTGCGGGAGATCC	16055 gp140 Flag/His-Tag QL	ATATTCGTCTCCTCGAGCTATCAGTGGTGATGGTGGTGGTGCTTGTCGTCGTCGTCCTTGTAGTCGCTGCCGCCTCCGCCGCTCTTGATGTACCACAGCCACTTGG
96ZM651 gp145 HA-Tag QL rev	ATATTCGTCTCCTCGAGCTATCATCAGGCGTAGTCGGGCACATCGTAGGGGTAGTAG	16055 SOSIP gp140 QL Rev	ATATATACTCGAGCGTCTCGTCGAGTCAGTCCAGGGCCAGCAGGTCC
96ZM651 gp140 QL rev	ATATTCGTCTCCTCGAGCTATCATCACTTGATGTACCACAGCCACTTGG	HXB2 QL for	ATAATACGTCTCGCTAGCATGCGGGTGAAAGAGAAGTACCAGCAC
96ZM651 gp140 Flag/His-Tag QL rev	ATATTCGTCTCCTCGAGCTATCAGTGGTGATGGTGGTGGTGCTTGTCGTCGTCGTCCTTGTAGTCGCTGCCGCCTCCGCCGCTCTTGATGTACCACAGCCACTTGG	HXB2 gp140 tag QL	ATATATCGTCTCCTCGAGTCAGTGGTGATGGTGGTGGTGCTTGTCGTCGTCGTCCTTGTAGTCGCTGCCGCCTCCGCCGCTCTTGATGTACCACAGCCAGTTGGTGATG
pWPXLd BG505 for BamHI	ACGGGATCCCGTCTCGCTAGCCGGACCGGCCACCATGAGAGTGATGGGCATCCAGCGGAACTGCCAGC	HXB2 gp145 QL rev	ATATTCGTCTCCTCGAGCTATCATCAGCTGTAGCCCTGCCG
pWPXLd BG505 rev EcoRI	CGGGAATTCGTCTCCTCGAGCCTGCAGGTCATCAGGCGTAGTCGGGCACATCGTAGGGGTAGTAGCCCTGCCGCACTCTGTGGATCACGCTCAGCAC	JRFL pWPXLd for	ACGGGATCCCGTCTCGCTAGCCGG
BG505 QL for	ATAATACGTCTCGCTAGCATGAGAGTGATGGGCATCCAGCG	JRFL pWPXLd gp145 rev	CGTCTCCTCGAGCCTGCAGG
BG505 gp145 QL rev	ATATTCGTCTCCTCGAGCTATCAGTAGCCCTGCCGCACTCTG	JRFL QL for	ATAATACGTCTCGCTAGCATGAGAGTGAAGGGCATCCGGAAG
BG505 gp140 Flag/His-Tag QL rev	ATATTCGTCTCCTCGAGCTATCAGTGGTGATGGTGGTGGTGCTTGTCGTCGTCGTCCTTGTAGTCGCTGCCGCCTCCGCCGCTGTCCAGGGCCAGCAGGTCC	JRFL gp140 Flag/His-tag QL rev	ATATATCGTCTCCTCGAGTCAGTGGTGATGGTGGTGGTGCTTGTCGTCGTCGTCCTTGTAGTCGCTGCCGCCTCCGCCGCTCTTGATGTACCACAGCCACTTGG
pWPXLD 16055 Pme1 for	GAGGTTTAAACTACGGGATCCCGTCTCGCTAGCCGGACCGGCCACCATGAGAGTGCGGGGCATCCTGCGGAACTACCAGCAG	JRFL gp145 QL rev	ATATTCGTCTCCTCGAGTCATCAGTAGCCCTGCCGCAC
Sequencing Primer			
pC-CMV-For	GTAGGCGTGTACGGTGGGAGG	16055 seq2 for	AGAGAGTGGGCAAGAAGCTGGC
pC-BGH-Rev	GCAACTAGAAGGCACAGTCGAGG	16055 seq3 for	CCTGCTGAAGGCCATCGAGG
pWPXLD seq for	ATCTCGACGGTATCGGTTAAC	HXB2 seq1 for	GATCATGGAAAAGGG
pWPXLd rev seq	GGGAATTCGTCTCCTCGAGCCTGC	HXB2 seq2 for	GCAGGCCCACTGCAAC
pQL13_seq_for	CATGGTCCTGCTGGAGTTCGTG	JRFL seq1 for	CGGGACGAGGTGCAGAAAG
96Zm651 seq1 for	ATCCCCATCCACTACTGCG	JRFL seq2 for	CCAGATCAGATGCAGCAGC
96ZM651 seq2 fwd	AGCACCAACGACAGCACC	BG505 seq1	GGACAAGAAGTTCAACGG
16055 seq1 for	GGTCAACACCACCAACGCC	BG505 seq2	AGAGAAGAGTCGTGGGACG

Appendix Table 2: Plasmid vectors used in this thesis

Description	Selection marker	Specification
pOG44	Amp^R	Mammalian vector for transient expression of the Flp recombinase[179] (Thermo Fisher Scientific, # V600520); used for generation of all stable cell lines of the FACS panning section
pWPXLd	Amp^R	Lentiviral vector (Addgene, #12258). Used for all gp145 envelope variants used in this thesis, including the 96ZM651 alanine library.
pcDNA3.1(+) QL	Amp^R	Derivative of pcDNA3.1(+) (Invitrogen/Thermo Fisher Scientific, # V790-20), including a CcdB cloning cassette inserted at the MCS; used for all gp140 envelope sequences, except the Env/V3 model library.
pQL13	Amp^R	Lentiviral vector including a CcdB cloning cassette and eGFP; used for generation of the gp145 Env/V3 stable cell lines of the FACS panning procedure and gp140 envelopes of the Env/V3 model library

Appendix Table 3: Alanine mutations potentially influencing envelope structure. A comparison of alanine mutants exhibiting loss or gain of signal effects in the alanine scans of sCD4, VRC01 and PG9 is given, including regions where those mutations are located, the exact amino acid position (HXB2 numbering) and 5F3 normalized fold changes compared to WT binding levels. Conservation in % was calculated based on an alignment of HIV-1 envelope sequences from 2013 available at the Los Alamos HIV Database (http://www.hiv.lanl.gov). Bold: statistically significant residues (FDR<0.05); underlined: contact residues to the respective screening reagent as determined by crystallography[20,42,117].

Region		AA Position	FC PG9	FC VRC01	FC sCD4	Conservation (%)	Region		AA Position	FC PG9	FC VRC01	FC sCD4	Conservation (%)
C1	β1	F93A	0.56	**0.57**	**0.03**	99.3	C2	loopC	L265A	0.66	**0.49**	0.58	97.2
C1	β1	N94A	**0.58**	0.55	**0.20**	95.3	C2	loopC	I270A	0.99	**0.49**	**0.30**	62.6
C1		W96A	0.59	**0.57**	**0.15**	99.1	C2	β10	I272A	0.90	**0.58**	**0.49**	94.6
C1		N98A	0.54	**0.58**	**0.14**	99.6	C2	loopD	L277A	0.90	0.55	0.46	42.1
C1	α1	M100A	**0.56**	**0.59**	**0.12**	99.6	C2	β11	I284A	0.81	**0.57**	**0.37**	98.8
C1	α1	E102A	**0.67**	0.76	**0.46**	63.5	C2	β11	I285A	0.65	**0.57**	**0.28**	96.9
C1	α1	Q103A	**0.66**	0.76	**0.52**	99.8	C2	β11	L288A	0.73	**0.35**	**0.04**	92.1
C1	α1	M104A	**0.66**	0.65	**0.42**	90.8	C3	β14	H330A	0.71	0.56	0.68	69.8
C1	α1	H105A	0.70	**0.57**	**0.36**	80.4	C3	β14	I333A	0.76	**0.42**	0.57	65.9
C1	α1	I109A	**0.71**	0.82	**0.59**	98.3	C3	α4	R335A	0.95	**0.57**	0.58	30.9
C1	α1	L111A	**2.27**	1.54	**0.51**	97.9	C3	α4	W338A	0.87	0.42	**0.24**	99.7
C2	β4	I215A	0.59	0.72	**0.14**	99.1	C3	α4	L342A	0.85	**0.40**	0.47	95.4
C2	β5	K227A	**0.60**	0.65	**0.29**	95.7	C3		L349A	**0.62**	0.37	0.69	97.9
C2	β5	N229A	**0.65**	0.71	**0.18**	78.6	C3		F353A	0.78	**0.30**	**0.22**	82.2
C2	loopA	F233A	0.83	**0.52**	**0.08**	99.1	C3	β14	F361A	0.93	**0.27**	**0.10**	98.7
C2	loopA	G235A	0.80	**0.53**	**0.05**	99.8	C3	α3	E370A	**0.86**	**0.53**	_0.63_	99.6
C2	β6	G237A	0.71	0.56	**0.23**	99.3	C3	α3	I371A	0.87	**_0.30_**	_0.48_	85.8
C2	β7	N241A	0.74	**0.59**	**0.41**	97.4	C3	β16	F376A	0.96	0.38	**0.27**	98.6
C2	β7	V242A	0.68	**0.60**	**0.26**	98.2	C3	β17	F383A	0.90	**0.41**	**0.27**	99.4
C2	β7	S243A	0.64	**0.58**	0.45	98.9	C3	α4	T387A	0.74	0.50	**0.34**	90.8
C2	β7	V245A	0.75	**0.65**	**0.47**	98.6	C3	α4	F391A	0.80	0.48	**0.12**	99.3
C2	β8	T248A	0.64	0.57	**0.13**	99.5	C4	β19	I414A	0.69	0.59	**0.20**	92.4
C2	β8	I251A	0.62	**0.53**	**0.35**	99.8	C4	β19	L416A	0.73	**0.55**	0.58	86.5
C2	loopB	K252A	0.68	**0.60**	0.52	66.2	C4	β19	R419A	0.58	**0.69**	1.54	82.7
C2	loopB	P253A	0.69	**0.59**	0.66	99.9	C4		P437A	0.70	**0.64**	0.52	95.7
C2	loopB	T257A	0.70	**0.53**	**0.40**	99.7	C4	β22	I449A	0.85	0.46	**0.27**	99.3
C2	loopB	Q258A	0.62	**0.33**	0.29	98.5	C4	β23	T450A	0.87	0.44	**0.35**	99.7
C2	loopB	L259A	0.75	**0.34**	**0.12**	99.0	C4	β23	L452A	0.94	**0.47**	**0.27**	80.3
C2	loopB	L260A	0.63	**0.39**	**0.13**	98.7	C4	β23	L454A	0.98	0.45	**0.22**	98.2
C2	β9	L261A	0.65	**0.39**	**0.08**	99.2	C5	β24	I467A	0.83	**_0.51_**	**0.32**	33.0
C2	β9	N262A	0.68	**0.37**	**0.16**	99.8	C5	β24	P470A	**0.69**	**0.34**	0.41	99.4
C2	β9	G263A	0.64	**0.54**	**0.10**	100.0	C5	α5	W479A	**0.56**	**0.44**	**0.24**	99.6
C2	loopC	S264A	0.56	**0.41**	**0.18**	99.0							

0 ▬▬▬ 1 ▬▬▬ 2

Appendix Table 4: Alanine scanning data of PG9, VRC01, sCD4 and 17b. 5F3 normalized fold changes are given for each variant relative to the 96ZM651 WT and statistically significant results (FDR<0.05) are bolded. Locations of the variants are indicated with i: inner domain, o: outer domain, b: bridging sheet, α: α helix, β: β sheet

Domain	Region	AA Position		Relative FC				Domain	Region	AA Position		Relative FC			
		96ZM651	HXB2	PG9	VRC01	sCD4	17b			96ZM651	HXB2	PG9	VRC01	sCD4	17b
		G31A	E32	1.05	1.01	**0.58**	0.74			D77A	D78	1.14	0.92	**0.46**	0.90
		N32A	K33	1.05	1.09	**0.51**	1.70			P78A	P79	1.00	0.90	**0.53**	0.64
		L33A	L34	1.08	1.04	0.63	1.13			N79A	N80	1.09	0.93	**0.70**	0.74
		W34A	W35	0.98	1.02	**0.44**	1.49			P80A	P81	1.21	0.90	0.43	0.75
		V35A	V36	1.23	1.33	**0.56**	1.12			Q81A	Q82	1.13	0.91	**0.46**	0.68
		T36A	T37	0.94	1.01	0.75	1.42			E82A	E83	0.75	0.69	0.31	**0.16**
		V37A	V38	1.05	1.15	**0.53**	1.26			I83A	V84	1.01	0.84	**0.75**	0.64
		Y38A	Y39	0.97	1.21	**0.53**	1.32			V84A	V85	0.75	0.88	**0.66**	0.53
		Y39A	Y40	1.07	1.42	**0.46**	1.41			L85A	L86	0.72	0.75	0.31	0.49
		G40A	G41	0.98	0.91	0.60	0.77			G86A	V87	0.95	1.10	**0.46**	0.88
		V41A	V42	1.08	0.98	**0.42**	0.69			N87A	N88	0.79	0.77	**0.48**	**0.50**
		P42A	P43	1.38	1.18	0.77	**0.21**			V88A	V89	1.02	0.83	0.43	0.95
		V43A	V44	1.55	1.02	**0.57**	0.78	i		T89A	T90	0.72	0.78	**0.50**	0.47
		W44A	W45	0.77	0.71	**0.39**	0.64	i		E90A	E91	0.89	0.86	**0.52**	**0.87**
		K45A	K46	1.03	0.87	0.60	0.76	i		N91A	N92	1.40	1.03	**0.61**	**0.33**
		E46A	E47	0.93	0.90	**0.46**	0.71	i	β1	F92A	F93	0.56	0.57	**0.03**	0.14
		K48A	T49	0.78	0.83	**0.37**	**0.48**	i		N93A	N94	0.58	0.55	**0.20**	**0.33**
		T49A	T50	0.94	0.87	0.56	1.07	i		M94A	M95	0.86	0.79	0.60	0.68
		T50A	T51	0.89	0.90	**0.46**	0.98	i		W95A	W96	0.59	0.57	0.15	0.23
		L51A	L52	0.86	0.79	**0.38**	**0.23**	i		K96A	K97	1.02	0.82	**0.49**	0.64
		F52A	F53	0.87	0.85	**0.45**	**0.12**	i		N97A	N98	0.54	0.58	0.14	0.15
		S55A	S56	1.03	0.89	**0.67**	0.90	i		D98A	D99	**0.80**	1.01	0.67	0.60
		D56A	D57	0.85	0.88	**0.59**	**0.66**	i	α1	M99A	M100	0.56	0.59	0.12	**0.54**
		K58A	K59	0.76	0.87	0.53	0.84	i		V100A	V101	0.86	0.83	**0.39**	0.43
		S59A	A60	0.71	0.90	0.71	0.65	i		D101A	E102	0.67	0.76	**0.46**	0.65
		Y60A	Y61	0.87	0.82	**0.47**	0.78	i		Q102A	Q103	0.66	**0.76**	**0.52**	**0.35**
		E61A	D62	0.94	0.88	0.66	0.80	i		M103A	M104	0.66	0.65	**0.42**	**0.28**
		K62A	T63	**0.78**	0.84	0.80	0.89	i		H104A	H105	0.70	**0.57**	**0.36**	**0.16**
		E63A	E64	0.73	0.82	0.58	**0.13**	i		E105A	E106	0.84	1.03	**0.52**	0.90
		V64A	V65	0.72	0.89	**0.53**	0.95	i		D106A	D107	0.63	0.76	**0.26**	**0.46**
		H65A	H66	0.81	0.80	**0.54**	0.46	i		I107A	I108	0.67	0.66	**0.46**	**0.29**
		N66A	N67	0.68	**0.72**	**0.38**	**0.33**	i		I108A	I109	**0.71**	0.82	**0.59**	0.74
		V67A	V68	0.75	0.80	**0.57**	**0.42**	i		S109A	S110	1.26	1.04	0.42	0.72
		W68A	W69	0.73	0.73	**0.40**	0.31	i		L110A	L111	**2.27**	1.54	**0.51**	**0.19**
		T70A	T71	0.84	0.85	**2.78**	0.60	i		W111A	W112	0.74	0.69	**0.16**	**0.14**
		H71A	H72	1.03	0.96	0.53	0.63	i		D112A	D113	1.15	0.94	**0.49**	1.07
		V74A	V75	1.02	0.98	**0.48**	0.13	i		Q113A	Q114	0.97	1.02	0.67	0.87
		P75A	P76	1.22	0.93	**0.53**	0.90	i		S114A	S115	1.22	1.15	0.64	0.62
		T76A	T77	1.13	0.94	0.56	0.71	i		L115A	L116	0.78	0.72	**0.53**	**0.20**

Domain	Region	AA Position 96ZM651	HXB2	Relative FC PG9	VRC01	sCD4	17b	Domain	Region	AA Position 96ZM651	HXB2	Relative FC PG9	VRC01	sCD4	17b
i	α1	K116A	K117	1.08	0.86	0.66	0.77	o		M158A	I154	0.59	0.84	**0.57**	0.67
i		P117A	P118	1.59	0.93	**0.33**	0.60	o	V1	K159A	K155	0.63	0.87	2.44	0.81
b		V119A	V120	1.49	0.94	**0.66**	0.38	o		N160A	N156	0.59	0.83	0.39	0.61
b		K120A	K121	1.25	0.82	**0.65**	**0.24**	o		S162A	S158	**0.42**	0.79	0.58	**0.45**
b	β2	L121A	L122	1.01	0.84	**0.54**	0.51	o		F163A	F159	0.49	0.74	**0.54**	**0.59**
b		T122A	T123	0.88	0.92	0.80	0.80	o		N164A	N160	0.44	1.11	0.61	1.20
b		P123A	P124	1.33	1.05	0.71	0.72	o		I165A	I161	1.01	0.80	**0.46**	**0.75**
		L124A	L125	**0.66**	0.89	0.72	1.39	o		T166A	S162	0.44	1.02	**0.54**	1.21
o		V126A	V127	**0.55**	0.91	0.57	1.34	o		T167A	T163	0.75	0.91	**0.47**	0.93
o		T127A	S128	1.05	1.01	0.76	1.25	o		E168A	S164	0.58	1.01	**0.47**	0.80
o		L128A	L129	**0.38**	0.82	**0.38**	1.09	o		L169A	I165	0.66	1.00	0.66	0.96
o		N129A	K130	1.60	1.06	0.76	0.91	o	V2	K170A	R166	0.82	0.98	0.64	0.74
o		T131A	T132	1.43	0.97	**0.55**	**0.63**	o		D171A	G167	0.85	0.92	0.55	**0.61**
o		E132A	D133	0.83	0.82	**0.46**	0.66	o		K172A	K168	**0.59**	0.86	0.78	0.84
o	V1	V133A	L134	0.79	0.98	0.62	0.65	o		K173A	V169	**0.51**	1.11	0.76	0.92
o		N134A	K135	1.12	1.02	0.74	0.87	o		K174A	Q170	0.88	0.98	0.77	0.90
o		V135A	N136	0.68	0.96	0.59	0.96	o		N175A	K171	1.38	1.08	**0.61**	0.69
o		T136A	D137	1.36	1.04	0.76	0.88	o		V176A	E172	0.96	1.11	0.72	1.06
o		R137A	T138	1.06	0.87	0.67	0.87	o		Y177A	Y173	0.78	0.97	NA	0.80
o		N138A	N139	1.22	1.00	0.82	0.89	o		L179A	F175	0.66	0.85	**0.34**	**0.56**
o		V139A	T140	1.27	1.14	0.70	0.82	o		F180A	F176	0.63	0.88	**0.36**	**0.67**
o		N140A	N141	1.29	1.09	**0.47**	0.79	o		Y181A	Y177	0.82	0.99	**0.65**	1.09
o		N141A	S142	1.06	1.03	0.62	0.98	o		K182A	K178	0.89	0.96	**0.54**	0.86
o		S142A	S143	1.16	1.04	0.65	1.01	o		L183A	L179	1.27	1.11	0.66	1.54
o		V143A	S144	1.14	1.05	0.56	0.99	o		D184A	D180	0.88	0.83	**0.49**	**0.58**
o		V144A	G145	1.00	1.05	**0.57**	1.13	o		I185A	I181	**0.57**	1.04	0.69	0.71
o		N145A	R146	0.98	1.06	0.78	0.82	o		V186A	I182	0.96	1.05	0.62	0.72
o		N146A	M147	1.09	0.99	0.82	1.02	o		S187A	P183	1.13	1.10	0.82	1.12
o		T147A	I148	0.95	0.97	0.69	1.17	o		L188A	I184	1.03	1.05	0.77	0.77
o		T148A	M149	1.08	1.06	0.62	1.01	o		N189A	D185	1.17	1.12	0.76	0.83
o		N149A	E150	1.12	1.08	0.71	1.00	o		E190A	N186	1.00	1.10	0.61	0.87
o		V150A	K151	0.74	0.94	**0.57**	0.90	o		T191A	-	1.15	1.13	0.87	0.82
o		N151A	-	1.53	1.26	0.56	1.27	o		D192A	-	0.99	1.10	0.67	0.88
o		N152A	-	1.32	1.24	**0.60**	0.97	o		D193A	-	0.98	0.98	**0.58**	**0.25**
o		S153A	-	1.74	1.33	**0.52**	1.04	o		S194A	-	0.89	1.04	**0.57**	0.78
o		M154A	-	0.72	1.03	**0.49**	0.89	o		E195A	-	0.99	0.94	**0.54**	0.80
o		N155A	-	1.16	1.35	**0.59**	1.43	o		T196A	-	0.95	0.99	0.67	1.12
o		G156A	G152	1.09	1.34	**0.50**	0.88	o		G197A	-	0.78	0.95	0.89	0.95
o		D157A	E153	0.79	0.93	2.42	0.88	o		N198A	-	0.91	1.05	**0.60**	0.97

Domain	Region	AA Position 96ZM651	HXB2	Relative FC PG9	VRC01	sCD4	17b	Domain	Region	AA Position 96ZM651	HXB2	PG9	VRC01	sCD4	17b
o		S199A	D187	0.88	1.09	0.67	1.07	i	loopA	G247A	G235	0.80	0.53	**0.05**	**0.11**
o		S200A	T188	0.94	1.08	0.64	0.96	i	loopA	T248A	T236	0.97	1.16	0.72	0.64
o		K201A	T189	0.89	0.95	0.63	1.01	i		G249A	G237	0.71	0.56	**0.23**	**0.19**
o		Y202A	S190	0.96	1.03	0.79	0.93	i	β6	P250A	P238	0.84	0.82	0.57	0.69
o	V2	Y203A	Y191	0.70	0.88	0.65	0.51	i		H252A	T240	0.92	0.80	0.84	0.54
o		R204A	K192	0.74	0.80	**0.55**	**0.54**	i		N253A	N241	0.74	**0.59**	**0.41**	**0.41**
o		L205A	L193	0.64	0.93	0.71	0.73	i	β7	V254A	V242	0.68	**0.60**	**0.26**	**0.31**
o		I206A	T194	0.86	1.06	0.66	1.01	i		S255A	S243	0.64	**0.58**	0.45	0.65
o		N207A	S195	0.94	0.86	**0.54**	0.83	i		T256A	T244	1.10	1.19	1.08	**0.17**
o		N209A	N197	0.65	0.96	0.47	0.63	i		V257A	V245	0.75	**0.65**	0.47	0.29
b		T210A	T198	1.34	0.91	0.71	0.74	i	β8	Q258A	Q246	0.89	0.81	**0.61**	0.75
b		S211A	S199	1.11	1.21	**0.56**	**0.57**	i		T260A	T248	0.64	0.57	**0.13**	**0.15**
b	β3	L213A	I201	1.05	0.86	0.84	0.69	i		H261A	H249	0.77	0.67	0.60	**0.34**
b		T214A	T202	1.04	0.87	0.94	0.65	i	β8	G262A	G250	0.76	0.79	0.92	**0.38**
b		Q215A	Q203	1.13	0.97	0.59	0.68	i		I263A	I251	0.62	**0.53**	**0.35**	**0.17**
i		P218A	P206	1.20	1.06	1.04	**0.55**	o		K264A	R252	0.68	**0.60**	0.52	**0.22**
i		K219A	K207	1.11	0.97	0.76	**0.28**	o		P265A	P253	0.69	**0.59**	0.66	**0.36**
i		V220A	V208	1.09	0.90	0.78	**0.51**	o		V266A	V254	0.65	**0.63**	0.70	**0.26**
i		S221A	S209	0.94	0.90	**0.46**	0.75	o		V267A	V255	0.80	0.87	0.95	**0.44**
i		F222A	F210	1.07	0.75	**0.43**	**0.16**	o	loopB	S268A	S256	0.77	0.66	0.62	**0.48**
i		D223A	E211	1.09	0.90	**0.55**	**0.24**	o		T269A	T257	0.70	**0.53**	**0.40**	**0.21**
i		P224A	P212	1.05	0.81	0.57	**0.21**	o		Q270A	Q258	0.62	**0.33**	0.29	**0.29**
i		I225A	I213	1.01	0.89	0.37	**0.18**	o		L271A	L259	0.75	**0.34**	**0.12**	**0.21**
i		P226A	P214	0.63	0.80	0.50	**0.14**	o		L272A	L260	0.63	**0.39**	**0.13**	**0.15**
i		I227A	I215	0.59	0.72	**0.14**	**0.28**	o		L273A	L261	0.65	**0.39**	**0.08**	**0.15**
i	β4	H228A	H216	0.85	0.91	**0.40**	**0.18**	o	β9	N274A	N262	0.68	**0.37**	**0.16**	**0.10**
i		Y229A	Y217	0.64	0.72	**0.21**	**0.12**	o		G275A	G263	0.64	**0.54**	**0.10**	**0.13**
i		P232A	P220	0.81	0.96	**0.53**	**0.21**	o		S276A	S264	0.56	**0.41**	**0.18**	**0.11**
i		G234A	G222	0.83	0.86	**0.20**	**0.55**	o		L277A	L265	0.66	**0.49**	0.58	**0.36**
i		Y235A	F223	0.80	0.79	**0.32**	**0.26**	o	loopC	E279A	E267	0.63	0.74	0.60	0.57
i	β5	I237A	I225	0.66	0.66	**0.05**	0.64	o		E280A	E268	0.74	0.84	0.83	0.95
i		L238A	L226	0.71	0.73	**0.24**	0.79	o		G281A	E269	0.78	0.80	0.99	1.20
i		K239A	K227	**0.60**	0.65	**0.29**	0.80	o		I282A	V270	0.99	**0.49**	**0.30**	**0.27**
i		N241A	N229	**0.65**	0.71	**0.18**	0.73	o		I283A	V271	0.98	0.76	0.77	**0.46**
i		N242A	N230	0.76	1.07	0.66	0.82	o	β10	I284A	I272	0.90	**0.58**	**0.49**	**0.30**
i		K243A	K231	1.07	0.80	0.78	0.61	o		R285A	R273	0.95	0.80	0.72	**0.49**
i	loopA	T244A	T2432	1.08	0.98	0.58	0.79	o		S286A	S274	1.03	0.84	0.85	0.69
i		F245A	F233	0.83	**0.52**	**0.08**	**0.18**	o		E287A	V275	0.98	0.85	0.62	0.57
i		N246A	N234	1.02	1.31	0.66	0.77	o	loopD	N288A	N276	0.86	0.87	**0.60**	0.80

Domain	Region	AA Position 96ZM651	HXB2	PG9	VRC01	sCD4	17b	Domain	Region	AA Position 96ZM651	HXB2	PG9	VRC01	sCD4	17b
o	loopD	L289A	F277	0.90	0.55	0.46	**0.30**	o	V3	T330A	T319	0.78	0.86	0.99	0.61
o		T290A	T278	0.98	**1.81**	**0.62**	0.83	o		G331A	I320	0.88	1.03	1.11	1.07
o		N291A	D279	0.90	**0.26**	0.74	**0.48**	o		D332A	G321	1.19	0.87	0.88	0.51
o		N292A	N280	1.01	**0.40**	0.92	0.82	o		I333A	K322	0.88	0.88	0.90	1.02
o		V293A	A281	1.07	1.21	1.08	0.96	o		I334A	I323	0.90	0.82	0.88	1.08
o		K294A	K282	0.99	0.71	0.68	0.56	o		G335A	G324	0.74	0.72	0.92	0.65
o		T295A	T283	0.72	**0.59**	**0.42**	**0.30**	o		D336A	N325	NA	NA	NA	1.58
o	β11	I296A	I284	0.81	**0.57**	**0.37**	**0.41**	o		I337A	M326	**0.76**	0.85	0.94	0.75
o		I297A	I285	0.65	**0.57**	**0.28**	**0.16**	o		R338A	R327	**0.75**	0.82	1.36	0.69
o		V298A	V286	0.88	**0.69**	0.68	0.44	o		Q339A	Q328	0.87	0.85	0.77	0.68
o		H299A	Q287	1.06	NA	0.64	0.81	o	β14	H341A	H330	0.71	0.56	0.68	0.60
o		L300A	L288	0.73	**0.35**	**0.04**	**0.13**	o		N343A	N332	0.76	**0.59**	0.85	0.54
o		N301A	N289	0.86	0.88	0.59	0.62	o		I344A	I333	0.76	**0.42**	0.57	**0.37**
o		R302A	T290	0.95	0.91	1.02	0.56	o		S345A	S334	0.71	**0.45**	0.65	0.65
o		S303A	S291	0.89	0.93	0.71	0.67	o	α2	R346A	R335	0.95	**0.57**	**0.58**	0.49
o	β12	I304A	V292	0.78	0.60	0.64	0.42	o		T347A	A336	1.08	0.90	1.03	1.52
o		E305A	E293	1.52	0.86	0.99	0.57	o		N348A	K337	1.01	0.86	0.93	1.43
o		I306A	I294	0.83	**0.54**	0.92	0.40	o		W349A	W338	0.87	0.42	**0.24**	0.66
o		V307A	N295	0.86	0.95	0.84	0.66	o		T350A	N339	0.94	**0.74**	0.78	1.84
o		V309A	T297	0.89	0.95	1.27	**0.40**	o		K351A	N340	0.95	0.80	0.93	0.61
o		R310A	R298	0.65	0.78	1.31	**0.38**	o		T352A	T341	1.00	0.71	0.96	0.86
o	V3	P311A	P299	0.68	0.76	1.20	0.53	o		L353A	L342	0.85	**0.40**	0.47	0.49
o		N312A	N300	0.67	0.90	0.93	0.91	o		R354A	K343	1.01	0.88	1.04	0.71
o		N313A	N301	0.69	**0.74**	1.02	0.84	o		E355A	Q344	0.94	0.78	1.06	1.50
o		N314A	N302	0.71	0.98	1.00	0.89	o		V356A	I345	1.06	**0.45**	0.83	0.54
o		T315A	T303	0.69	0.73	**0.75**	1.13	o		R357A	A346	0.92	0.97	1.16	0.71
o		R316A	R304	0.59	0.80	0.91	0.84	o		N358A	S347	1.04	0.82	1.17	1.14
o		Q317A	K305	0.68	0.86	1.03	0.86	o		K359A	K348	0.93	0.87	0.91	1.17
o		S318A	R306	0.80	0.95	1.11	1.05	o		L360A	L349	**0.62**	0.37	0.69	1.33
o		I319A	I307	1.02	0.87	0.88	1.29	o		R361A	R350	0.88	0.87	1.14	1.38
o		R320A	R308	1.19	0.94	0.99	0.86	o		E362A	E351	1.45	0.85	**2.88**	1.54
o		I321A	I309	0.96	0.91	0.87	0.85	o		H363A	Q352	0.90	0.69	1.25	0.91
o		G322A	Q310	1.63	1.08	0.95	0.76	o		F364A	F353	0.78	**0.30**	**0.22**	0.65
o		P323A	R311	1.02	0.85	1.05	0.71	o		P365A	G354	0.89	**0.61**	0.88	0.74
o		G324A	G312	0.94	0.85	1.11	0.49	o		N366A	N355	0.93	0.68	1.08	0.95
o		Q325A	P313	0.98	0.85	0.88	0.86	o		K367A	K357	0.81	0.69	1.11	0.58
o		T326A	R315	0.90	**0.81**	0.86	1.07	o		N368A	T358	0.90	0.83	1.20	0.78
o		F327A	A316	0.85	0.97	0.83	0.64	o	β14	I369A	I359	0.80	**0.42**	0.80	**0.28**
o		Y328A	F317	0.85	0.96	1.07	0.63	o		T370A	T360	0.88	1.07	0.97	1.14

Domain	Region	AA Position 96ZM651	HXB2	Relative FC PG9	VRC01	sCD4	17b	Domain	Region	AA Position 96ZM651	HXB2	Relative FC PG9	VRC01	sCD4	17b
o	β14	F371A	F361	0.93	**0.27**	**0.10**	**0.25**	o	V4	G412A	G410	0.76	1.03	1.08	0.86
o		K372A	K362	0.84	1.02	1.09	1.09	o		T413A	D412	0.76	0.99	1.24	0.41
o		P373A	Q363	0.84	0.63	0.75	0.64	o		P414A	T413	**0.75**	0.99	1.00	1.08
o		S374A	S364	0.74	0.64	1.12	0.95	o		I415A	I414	0.69	0.59	**0.20**	**0.20**
o		S375A	S365	0.81	0.91	0.95	0.96	o		T416A	T415	0.72	0.67	0.99	0.50
o		G376A	G366	0.90	**0.36**	0.64	0.81	o	β19	L417A	L416	0.73	**0.55**	0.58	**0.33**
o	β15	G377A	G367	0.85	**0.29**	**0.27**	1.23	o		P418A	P417	0.64	**0.63**	0.76	0.45
o		D378A	D368	0.72	**0.49**	**0.26**	0.88	o		R420A	R419	0.58	**0.69**	1.54	0.26
o		L379A	P369	0.81	**0.23**	0.91	0.96	o		I421A	I420	1.00	**0.63**	0.69	NA
o	α3	E380A	E370	**0.86**	**0.53**	0.63	0.78	o		R422A	K421	0.82	0.70	0.90	0.36
o		I381A	I371	0.87	**0.30**	**0.48**	0.75	b		Q423A	Q422	0.98	0.78	1.07	1.20
o		T382A	V372	1.17	0.81	1.05	0.75	b	β20	I424A	I423	0.91	0.80	1.86	**0.25**
o		T383A	T373	1.19	0.72	0.82	0.63	b		I425A	I424	0.86	0.92	0.95	0.19
o		H384A	H374	1.19	**0.71**	0.91	NA	b		N426A	N425	0.96	0.75	**0.45**	0.53
o		S385A	S375	1.20	0.82	1.02	1.07	b		M427A	M426	0.93	0.83	0.91	0.95
o	β16	F386A	F376	0.96	0.38	**0.27**	0.94	b		W428A	W427	0.85	0.81	**0.07**	**0.13**
o		N387A	N377	1.08	0.64	0.97	1.03	b		Q429A	Q428	1.24	0.84	0.88	0.28
o		R389A	G379	1.19	0.88	0.82	0.40	b		E430A	K429	1.33	1.04	1.14	NA
o		G390A	G380	0.89	**0.54**	0.63	0.60	b	β21	V431A	V430	NA	NA	NA	NA
o		E391A	E381	1.23	0.94	1.20	0.60	b		G432A	G431	0.98	0.76	0.88	NA
o	β17	F392A	F382	0.97	**0.55**	0.73	**0.26**	b		R433A	K432	0.66	0.60	1.36	0.61
o		F393A	F383	0.90	**0.41**	**0.27**	0.45	b		M435A	M434	0.72	0.62	1.06	0.84
o		Y394A	Y384	1.02	**0.49**	1.15	0.60	b		Y436A	Y435	0.64	0.58	0.90	0.89
o		N396A	N386	0.95	**0.53**	0.62	0.85	o		P438A	P437	0.70	**0.64**	0.52	0.87
o	α4	T397A	S387	0.74	0.50	**0.34**	0.57	o		P439A	P438	0.75	0.74	1.27	0.85
o		S398A	T388	0.92	0.75	0.84	0.97	o		I440A	I439	0.76	0.74	1.21	1.01
o		G399A	Q389	0.91	1.01	1.08	**0.24**	o		E441A	S440	0.87	0.83	1.32	0.88
o		L400A	L390	0.90	0.76	0.85	1.07	o		G442A	G441	**0.69**	0.68	0.94	0.84
o		F401A	F391	0.80	0.48	**0.12**	**0.14**	o		N443A	Q442	0.89	1.13	1.15	0.42
o		S402A	N392	0.87	0.82	1.15	0.96	o		I444A	I443	0.73	0.65	0.85	NA
o	β18	I403A	S393	0.85	0.82	1.22	0.89	o	β22	K447A	S446	0.80	0.64	1.20	**0.35**
o		N404A	T394	0.88	0.93	1.14	0.77	o		S448A	S447	0.62	**0.46**	0.75	**0.34**
o		Y405A	W395	0.87	0.81	1.27	1.40	o		D449A	N448	1.09	1.08	0.80	0.64
o		T406A	F396	0.82	0.96	1.15	0.75	o		I450A	I449	0.85	**0.46**	**0.27**	**0.43**
o	V4	E407A	E403	1.17	1.04	1.24	0.76	o		T451A	T450	0.87	0.44	**0.35**	0.57
o		N408A	N406	0.85	0.93	1.24	0.86	o		G452A	G451	1.31	0.68	0.75	0.87
o		N409A	N407	0.78	1.03	1.09	0.84	o	β23	L453A	L452	0.94	**0.47**	**0.27**	0.54
o		T410A	T408	0.76	0.94	1.14	0.50	o		L454A	L453	0.92	**0.68**	0.77	**0.41**
o		D411A	E409	0.77	0.92	1.01	1.16	o		L455A	L454	0.98	0.45	**0.22**	**0.27**

| Domain | Region | AA Position | | Relative FC | | | | Domain | Region | AA Position | | Relative FC | | | |
		96ZM651	HXB2	PG9	VRC01	sCD4	17b			96ZM651	HXB2	PG9	VRC01	sCD4	17b
o	β23	V456A	T455	1.06	1.19	1.20	**0.23**	i	α5	S486A	S481	1.07	0.95	**0.12**	0.55
o		R457A	R456	0.70	**0.33**	**0.16**	**0.22**	i		E487A	E482	0.98	0.68	0.57	**0.17**
o		D458A	D457	0.92	**0.45**	**0.09**	**0.28**	i		L488A	L483	0.96	0.73	0.75	**0.15**
o	V5	G459A	G458	0.93	**0.66**	0.69	**0.31**	i		Y489A	Y484	1.06	0.66	0.91	0.30
o		G460A	G459	0.87	**0.51**	1.03	0.51	i		K490A	K485	0.97	0.79	1.02	0.40
o		S461A	N460	1.13	0.83	0.96	0.70	i		Y491A	Y486	0.91	0.56	1.05	0.12
o		T462A	S461	0.96	0.94	0.80	0.58	i		K492A	K487	0.88	0.81	0.66	0.25
o		N463A	NA	0.88	1.21	1.16	1.00	i	β25	V493A	V488	1.03	0.77	**0.61**	**0.17**
o		D464A	NA	0.89	0.64	1.03	1.05	i		V494A	V489	0.93	0.75	0.51	**0.23**
o		S465A	NA	0.92	1.29	1.20	1.06	i		E495A	K490	1.05	0.91	1.01	**0.25**
o		T466A	NA	0.94	0.94	2.21	1.07	i		I496A	I491	1.00	0.90	0.78	**0.23**
o		N467A	N462	0.86	0.68	1.62	1.09	i		K497A	E492	1.10	1.08	1.06	**0.34**
o		N468A	N463	0.97	0.79	1.51	1.08	i		P498A	P493	0.97	1.13	1.36	0.51
o		N469A	E464	0.96	0.90	1.22	1.05	i		L499A	L494	0.91	0.98	0.91	**0.38**
o	β24	T470A	S465	0.82	0.89	1.63	0.73	i		G500A	G495	0.96	0.99	1.18	0.40
o		E471A	E466	0.82	**0.54**	0.87	0.66	i		I501A	V496	0.92	0.86	0.90	**0.31**
o		I472A	I467	0.83	**0.51**	**0.32**	NA	i		P503A	P498	0.97	0.94	0.92	**0.30**
o		F473A	F468	NA	NA	**0.01**	1.11	i		T504A	T499	0.94	0.95	0.79	NA
o		R474A	R469	0.82	**0.31**	**0.28**	0.97	i		E505A	K500	0.78	1.01	0.83	**0.37**
o		P475A	P470	**0.69**	**0.34**	**0.41**	0.83	i		K507A	K502	0.82	1.00	0.89	NA
o		G477A	G472	0.72	**0.42**	**0.12**	0.91	i		R508A	R503	0.90	0.96	1.20	0.45
o		G478A	G473	0.84	0.60	**0.30**	0.47	i		R509A	R504	0.86	1.01	1.13	0.68
o		D479A	D474	0.78	**0.52**	0.57	**0.19**	i		V510A	V505	0.85	0.99	0.94	0.55
i	α5	M480A	M475	0.82	0.78	0.81	**0.12**	i		V511A	V506	0.64	0.95	0.84	0.68
i		R481A	R476	0.75	0.55	0.50	**0.12**	i		E512A	Q507	0.93	1.02	1.29	0.60
i		D482A	D477	0.93	0.55	0.88	0.26	i		R513A	R508	0.77	1.03	1.14	NA
i		N483A	N478	0.86	0.57	0.80	0.40	i		E514A	E509	**0.67**	0.84	1.23	0.27
i		W484A	W479	**0.56**	**0.44**	**0.24**	0.20	i		K515A	K510	0.78	1.04	1.46	NA
i		R485A	R480	0.66	0.76	0.88	**0.29**	i		S516A	R511	0.81	1.01	1.27	0.63

Clade	Isolate	PG9 FC S158T
A	BG505	0.95 ± 0.08
B	HXB2	0.58 ± 0.03
B	JRFL	0.87 ± 0.08
C	16055	0.95 ± 0.08
C	96ZM651	1.83 ± 0.19

Appendix Table 5: Transfer of mutation S158T to different isolates. S158T was analyzed in gp145 variants of the indicated isolates via flow cytometry (n=3). 5F3 normalized fold changes of PG9 binding levels relative to the respective WT are shown. The gain of signal effect observed for 96ZM651 was not transferable.

| Variant	T_M
96ZM651 WT	58.4 ± 1.8
96ZM651 L111A	58.9 ± 0.7
96ZM651 T278A	56.6 ± 0.1
96ZM651 W427A	59.1 ± 0.5
96ZM651 L111A/T278A	57.5 ± 0.1
96ZM651 L111A/T278H	61.1 ± 0.7
96ZM651 L111A/T278A/W427A	59.45 ± 0.8

Appendix Table 6: NanoDSF analysis of 96ZM651 single, double and triple mutants. (n=2). Insertion of L111A, T278A/H, W427A or combinations of these mutations did not significantly alter trimer thermostability. T_M: melting temperature.

Appendix Table 7: K$_D$ values of gp140 WT, L111A/T278H (LT – EDC/NHS) and L111A/T278H cross-linked with EDC/NHS (LT + EDC/NHS) from different isolates to a panel of Envelope antibodies targeting different epitopes. K$_D$ values affiliate with Figure 38 of the main results section. n=2.

| | non- or weakly neutralizing | | | | CD4 binding site | | | | V3 | | | | Apex | | | | gp120/gp41 Interface | |
| | 17b | | F105 | | HJ16 | | VRC01 | | HGN194 | | PGT121 | | PG9 | | PGT145 | | PGT151 | |
	CAPTURE	DIRECT	CAPTURE	DIRECT	CAPTURE	DIRECT	CAPTURE	DIRECT	CAPTURE	DIRECT	CAPTURE	DIRECT	CAPTURE	DIRECT	CAPTURE	DIRECT	CAPTURE	DIRECT
96ZM651																		
WT	0.78 ± 0.27	1.67 ± 0.29	0.22 ± 0.03	0.34 ± 0.04	0.05 ± 0.06	0.07 ± 0.00	3.87 ± 1.08	4.13 ± 0.65	0.24 ± 0.05	0.26 ± 0.13	1.09 ± 0.02	0.17 ± 0.08	49.68 ± 7.57	25.80 ± 3.25	ND	ND	ND	ND
LT - EDC/NHS	90.21 ± 18.94	97.24 ± 8.44	0.15 ± 0.03	0.30 ± 0.00	18.81 ± 6.72	31.96 ± 6.17	0.25 ± 0.11	0.30 ± 0.08	0.24 ± 0.01	0.25 ± 0.15	0.43 ± 0.09	0.08 ± 0.02	7.87 ± 1.27	5.21 ± 1.26	ND	ND	ND	ND
LT + EDC/NHS	115.00 ± 7.50	117.35 ± 4.74	0.49 ± 0.04	0.97 ± 0.22	31.52 ± 14.42	47.82 ± 18.30	0.35 ± 0.12	0.53 ± 0.04	0.25 ± 0.02	0.21 ± 0.07	0.68 ± 0.23	0.11 ± 0.02	13.53 ± 3.17	6.50 ± 0.61	ND	ND	ND	ND
16O55																		
WT	0.34 ± 0.06	0.65 ± 0.04	0.40 ± 0.24	0.33 ± 0.34	0.06 ± 0.01	0.07 ± 0.02	1.13 ± 0.73	2.74 ± 0.35	0.10 ± 0.01	0.11 ± 0.03	0.27 ± 0.23	0.19 ± 0.13	0.32 ± 0.18	0.45 ± 0.09	ND	ND	41.90 ± 59.25	63.40 ± 20.73
LT - EDC/NHS	7.90 ± 6.98	38.85 ± 5.89	0.65 ± 0.38	0.63 ± 0.40	7.08 ± 0.14	3.57 ± 2.30	0.25 ± 0.12	0.66 ± 0.16	0.10 ± 0.02	0.09 ± 0.06	0.18 ± 0.13	0.11 ± 0.06	0.16 ± 0.04	0.40 ± 0.08	ND	ND	48.11 ± 61.82	64.96 ± 28.92
LT + EDC/NHS	7.54 ± 7.68	50.93 ± 6.87	5.29 ± 0.12	3.19 ± 2.84	14.25 ± 1.58	8.11 ± 4.88	0.40 ± 0.03	1.21 ± 0.67	0.11 ± 0.03	0.12 ± 0.07	1.29 ± 0.68	0.32 ± 0.18	0.35 ± 0.19	0.89 ± 0.17	ND	ND	62.68 ± 77.24	82.65 ± 38.83
16055 SOSIP																		
WT	0.23 ± 0.08	1.22 ± 0.19	1.41 ± 0.69	0.94 ± 0.72	0.12 ± 0.01	0.10 ± 0.02	1.63 ± 1.20	3.66 ± 0.02	0.10 ± 0.03	0.10 ± 0.06	0.27 ± 0.36	0.26 ± 0.10	0.62 ± 0.50	1.04 ± 0.33	0.06 ± 0.07	0.00 ± 0.00	1.33 ± 1.87	2.21 ± 3.08
LT - EDC/NHS	2.05 ± 1.64	23.00 ± 1.29	3.94 ± 1.47	1.32 ± 1.12	2.57 ± 0.49	1.42 ± 0.79	0.45 ± 0.29	0.86 ± 0.00	0.12 ± 0.05	0.09 ± 0.06	0.07 ± 0.17	0.13 ± 0.08	0.33 ± 0.19	0.68 ± 0.28	0.05 ± 0.01	0.00 ± 0.00	1.06 ± 1.48	1.80 ± 2.23
LT + EDC/NHS	ND	12.26 ± 5.37	47.94 ± 39.53	1.68 ± 1.98	4.63 ± 0.09	3.59 ± 2.83	0.54 ± 0.14	1.57 ± 0.33	0.11 ± 0.05	0.08 ± 0.04	0.27 ± 0.52	0.28 ± 0.15	0.54 ± 0.34	1.57 ± 0.57	0.05 ± 0.01	0.06 ± 0.05	6.93 ± 9.76	4.01 ± ND
BG505																		
WT	4.97 ± 1.30	6.77 ± 3.30	0.39 ± 0.21	1.08 ± 0.51	0.27 ± 0.21	0.44 ± 0.42	0.12 ± 0.15	0.61 ± 0.51	0.06 ± 0.00	0.12 ± 0.08	0.58 ± 0.19	0.41 ± 0.08	1.79 ± 1.16	2.11 ± 0.09	ND	ND	90.98 ± ND	ND
LT - EDC/NHS	82.43 ± 4.07	101.67 ± 38.52	2.82 ± 2.41	1.78 ± 0.67	ND	ND	0.05 ± 0.10	0.19 ± 0.00	0.07 ± 0.01	0.13 ± 0.09	0.55 ± 0.25	0.35 ± 0.11	0.86 ± 0.76	1.08 ± 0.30	ND	ND	95.02 ± ND	ND
LT + EDC/NHS	143.31 ± 66.60	125.18 ± 55.32	10.58 ± 10.41	6.49 ± 2.34	ND	ND	0.07 ± 0.15	0.18 ± 0.11	0.07 ± 0.02	0.13 ± 0.05	0.70 ± 0.28	0.51 ± 0.19	1.37 ± 1.29	1.89 ± 0.08	ND	ND	ND	ND
BG505 SOSIP																		
WT	8.98 ± 0.28	3.46 ± 1.31	2.48 ± 0.01	0.63 ± 0.57	0.72 ± 0.39	0.17 ± 0.02	0.25 ± 0.07	0.35 ± 0.16	0.09 ± 0.02	0.13 ± 0.03	0.73 ± 0.11	0.26 ± 0.21	3.79 ± 3.99	1.64 ± 0.49	0.31 ± 0.24	0.52 ± 0.28	0.95 ± 0.22	2.65 ± 3.24
LT - EDC/NHS	141.10 ± 60.39	83.48 ± 23.26	3.10 ± 0.41	1.11 ± 1.04	62.38 ± ND	67.50 ± 95.18	0.04 ± 0.00	0.15 ± 0.05	0.10 ± 0.05	0.12 ± 0.11	0.80 ± 0.23	0.33 ± 0.18	3.34 ± 3.43	1.06 ± 0.81	0.55 ± 0.47	0.45 ± 0.43	0.99 ± 0.12	1.73 ± 0.50
LT + EDC/NHS	105.31 ± 76.36	92.94 ± 34.74	10.19 ± 4.79	10.18 ± 7.45	ND	156.14 ± 220.27	0.06 ± 0.03	0.23 ± 0.06	0.11 ± 0.05	0.12 ± 0.08	1.06 ± 0.15	0.63 ± 0.15	3.39 ± 2.95	3.52 ± 1.83	0.46 ± 0.45	0.69 ± 0.76	1.43 ± 0.02	4.12 ± 3.46

K$_D$ in nM:

| 0.005 | 0.01 | 0.1 | 0.25 | 0.5 | 1 | 2 | 4 | 8 | 15 | 31.25 | 75 | 100 | 125 | 150 | 175 | 200 |

WT   L111A

rel. MFI (PG9/5F3)

PG9 in nM
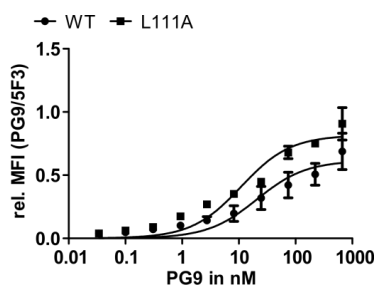
**Appendix Figure 1: PG9 FACS titration of 96ZM651 gp145 WT and L111A.**
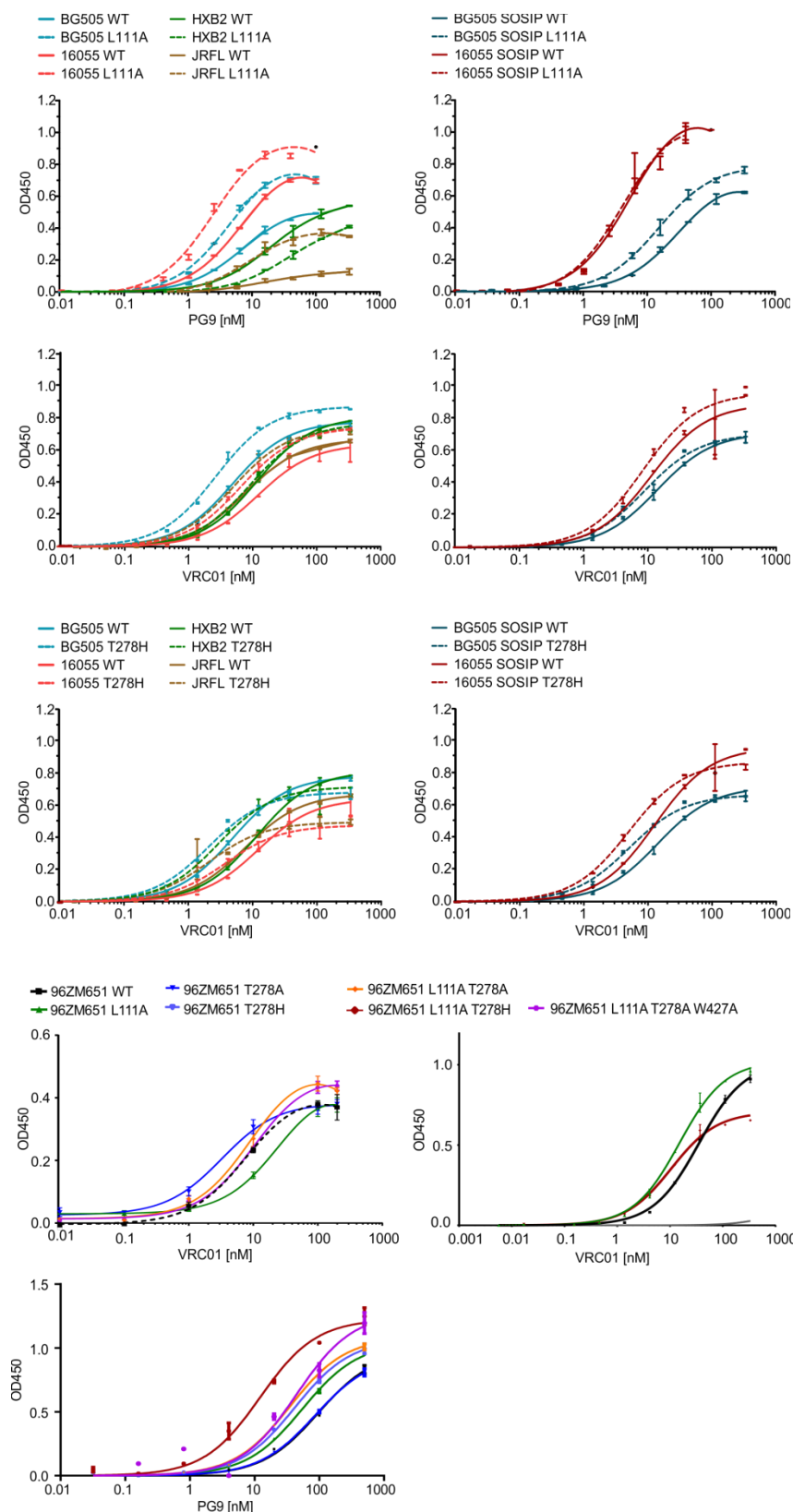
Appendix Figure 2: ELISA titration of soluble gp140 trimers from different isolates. Secreted gp140 trimers were purified from the supernatant of HEK293-F suspension cells using lectin affinity chromatography and subsequent size exclusion. Trimers were compared for VRC01 and PG9 binding in relation to the respective WT. Curves were analyzed with GraphPad Prism 5.0, non-linear regression (hyperbolar one-site binding. Data were generated by Christina Schmalzl and Matthias Glögl during my experimental supervision and by me.
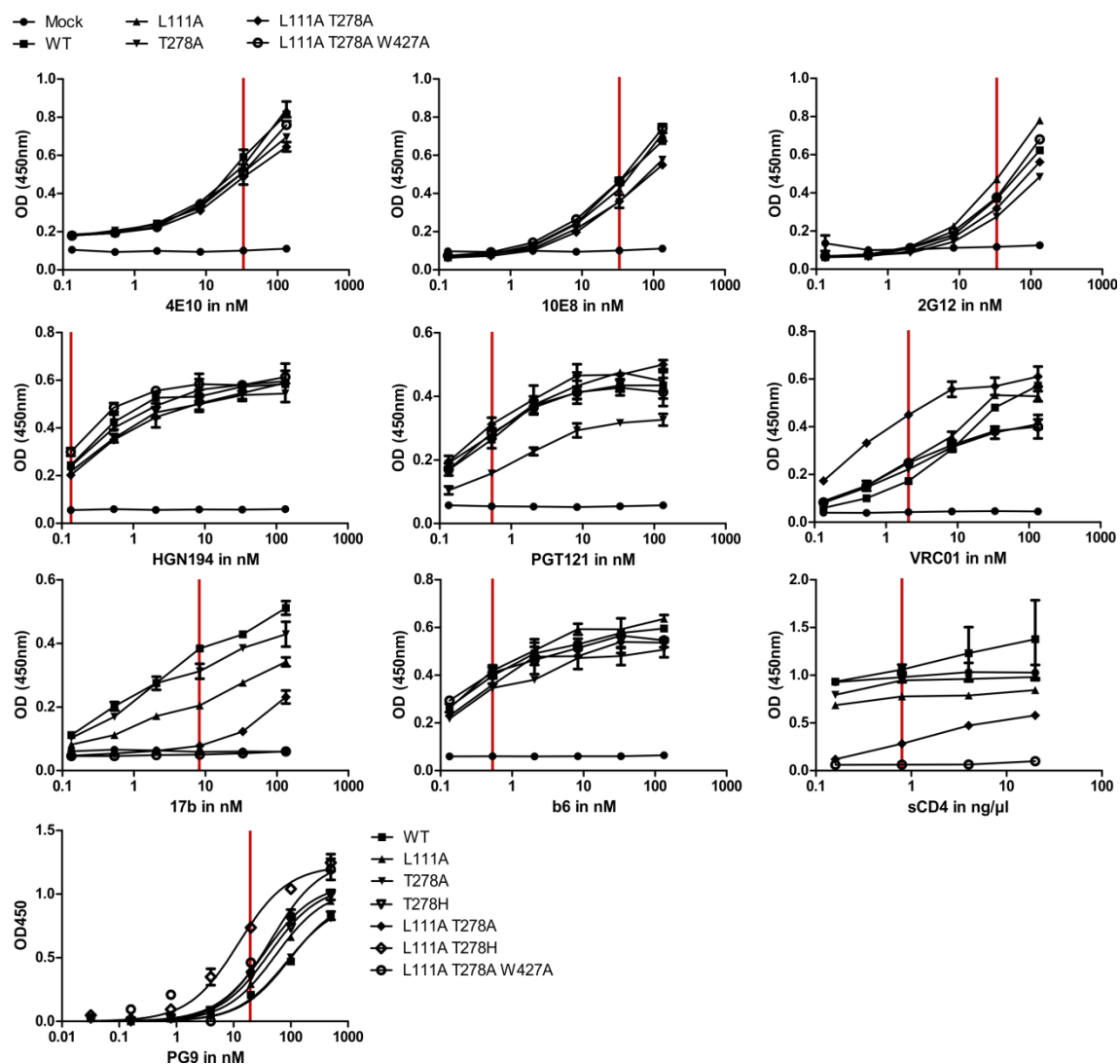
Appendix Figure 3: ELISA titration of 96ZM651 gp140 variants with a panel of different antibodies and sCD4. 96ZM651 gp140 trimers were titrated to analyze the influence of mutations L111A, T278A/H, W427A and their combinations on Env antigenicity. One concentration located in the linear increase of the sigmoidal binding curve was chosen (highlighted in red) to compare binding fold changes in the heat map formats of the main results section (Tables 7, 9 and 12).
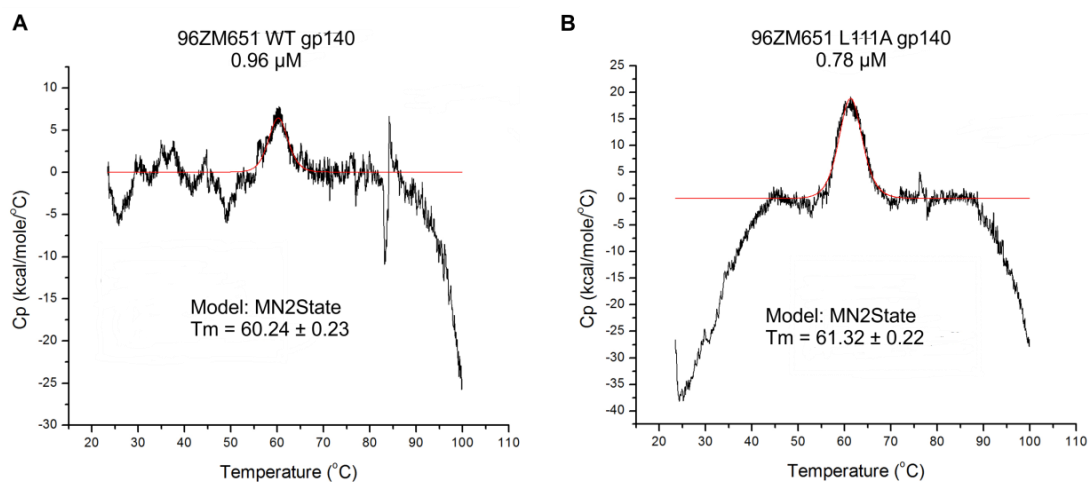
Appendix Figure 4: DSC analysis of 96ZM651 gp140 wildtype and mutant L111A. Although comparable protein concentrations were administered, 96ZM651 WT (A) depicted only a weak signal intensity and very inhomogeneous raw data. For mutant L111A (B), higher signal intensity and a more homogeneous melting curve could be obained. No significant difference was observed for WT and L111A thermostability. Black: raw data; red: fitted data used for the calculation of melting temperatures (Tm).
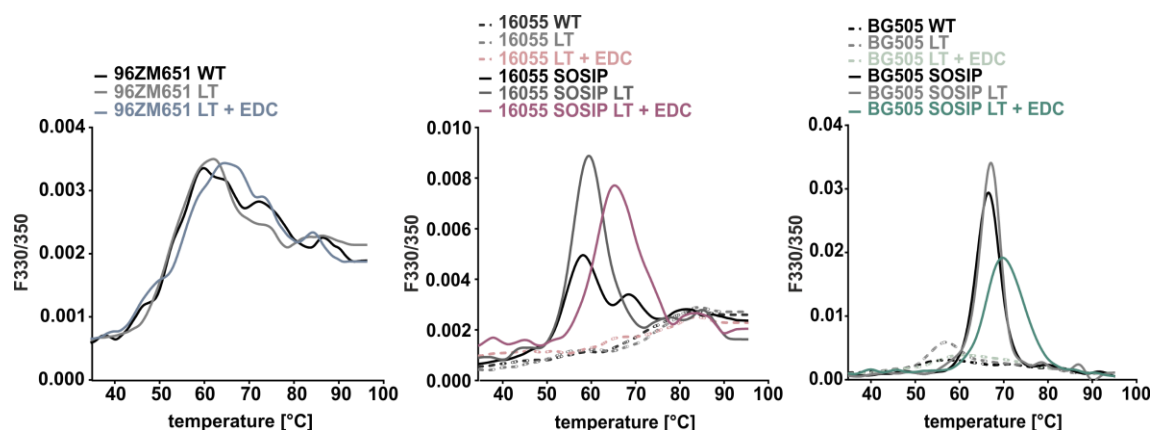

Appendix Figure 5: NanoDSF analysis of cross-linked envelope trimers. Thermostability of EDC/NHS cross-linked L111A/T278H variants (LT+EDC) was analyzed by nanoDSF in comparison to unmodified WTs and L111A/T278H (LT) variants (n=2). Note that only signal intensities of 16055 SOSIP and BG505 SOSIP allowed reliable analysis of melting temperatures. Signals for 96ZM651, 16055 and BG505 were low supposedly due to inhomogeneous folding caused by furin cleavage site mutation.
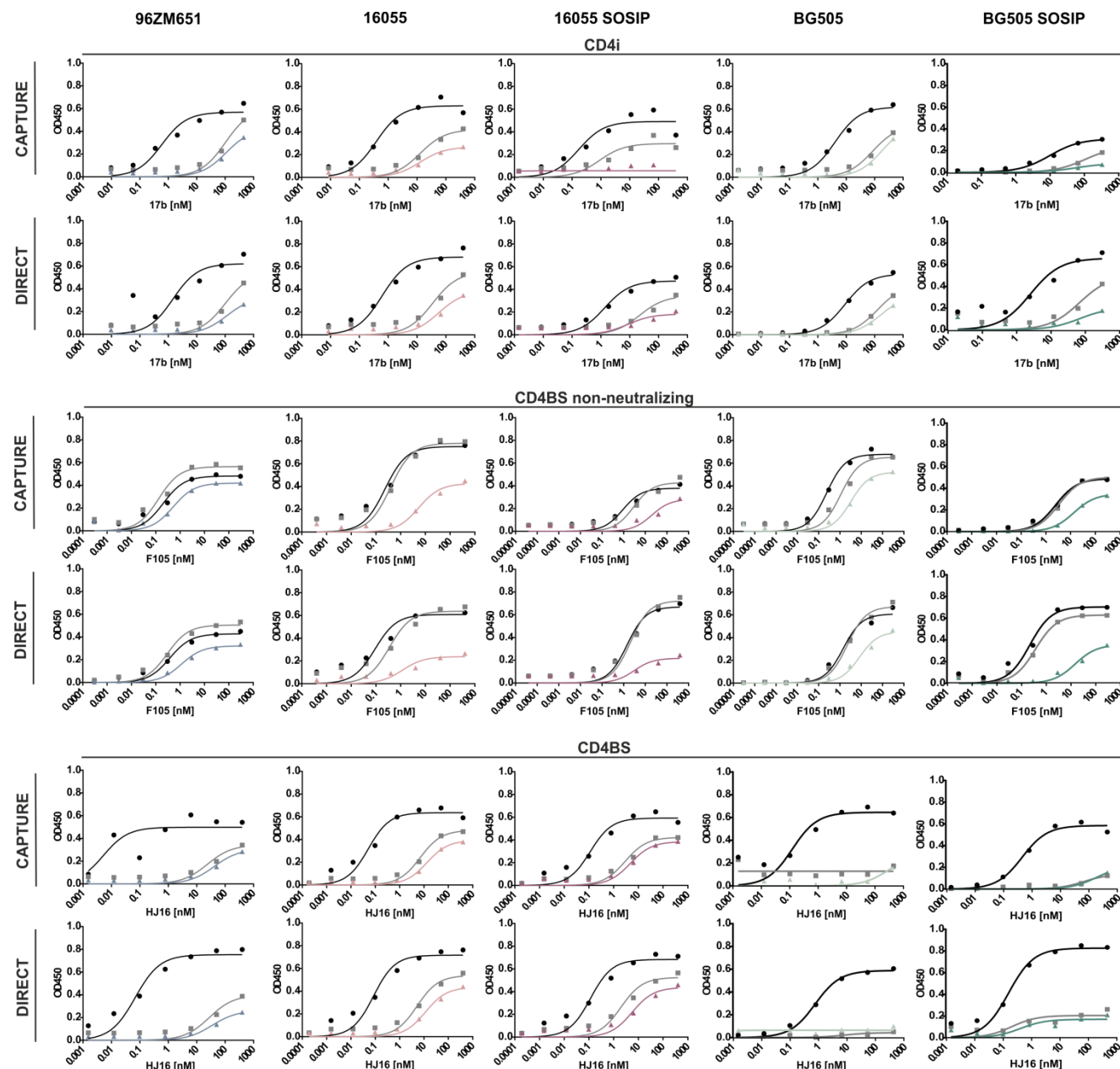
Appendix Figure 6A: Antigenicity profile of EDC/NHS cross-linked soluble envelope trimers. Antibody binding to WT, L111A/T278H and EDC/NHS cross-linked L111A/T278H gp140 trimers was analyzed by lectin capture (preserves trimer integrity) and direct ELISA (mildly denaturing conditions supporting trimer opening/unfolding). A panel of different antibodies was used including the non-neutralizing mAbs 17b (CD4i) and F105 (CD4BS), as well as bnAb HJ16 (CD4BS). One of two biological replicates is shown. The second dataset, which depicted the same results, is shown in Figure 37A. L111A/T278H resulted in decreased 17b recognition, which could further be enhanced by EDC/NHS cross-linking. Also, cross-linking diminished F105 binding. HJ16 recognition was impeded due to T278H, but was not further influenced by cross-linking.
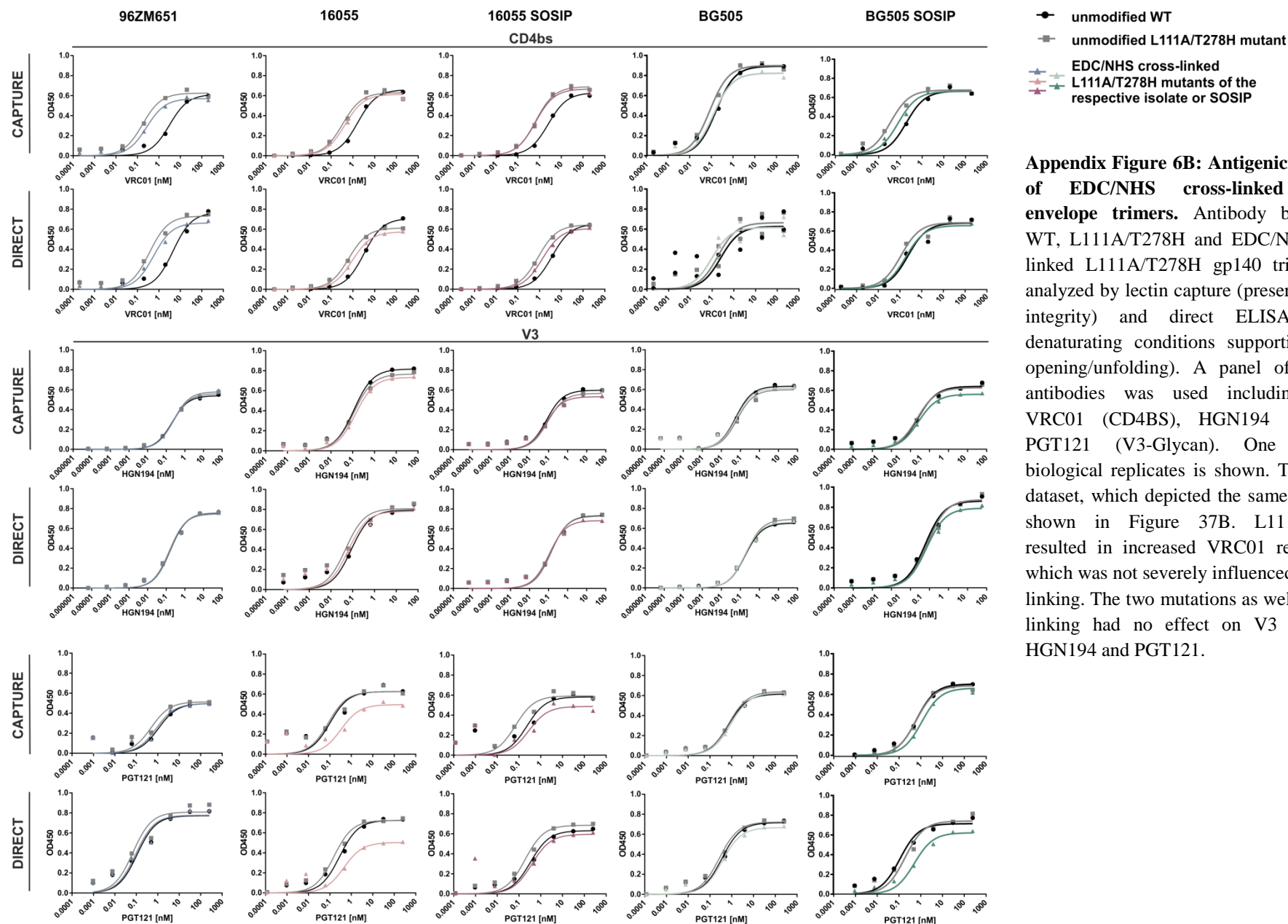
Appendix Figure 6B: Antigenicity profile of EDC/NHS cross-linked soluble envelope trimers. Antibody binding to WT, L111A/T278H and EDC/NHS cross-linked L111A/T278H gp140 trimers was analyzed by lectin capture (preserves trimer integrity) and direct ELISA (mildly denaturing conditions supporting trimer opening/unfolding). A panel of different antibodies was used including bnAbs VRC01 (CD4BS), HGN194 (V3) and PGT121 (V3-Glycan). One of two biological replicates is shown. The second dataset, which depicted the same results, is shown in Figure 37B. L111A/T278H resulted in increased VRC01 recognition, which was not severely influenced by cross-linking. The two mutations as well as cross-linking had no effect on V3 antibodies HGN194 and PGT121.
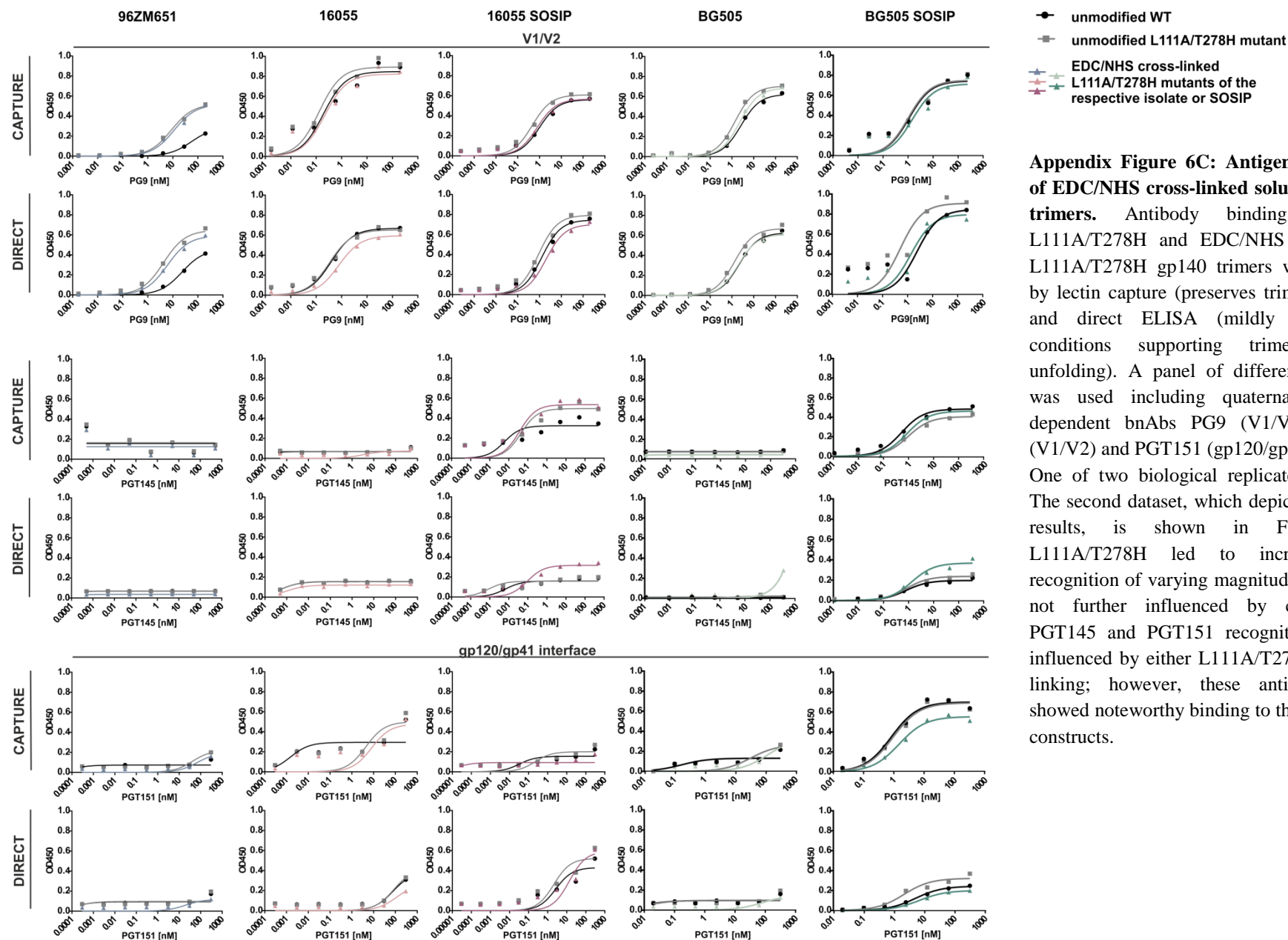
Appendix Figure 6C: Antigenicity profile of EDC/NHS cross-linked soluble envelope trimers. Antibody binding to WT, L111A/T278H and EDC/NHS cross-linked L111A/T278H gp140 trimers was analyzed by lectin capture (preserves trimer integrity) and direct ELISA (mildly denaturing conditions supporting trimer opening/ unfolding). A panel of different antibodies was used including quaternary structure dependent bnAbs PG9 (V1/V2), PGT145 (V1/V2) and PGT151 (gp120/gp41 interface). One of two biological replicates is shown. The second dataset, which depicted the same results, is shown in Figure 37C. L111A/T278H led to increased PG9 recognition of varying magnitude, which was not further influenced by cross-linking. PGT145 and PGT151 recognition was not influenced by either L111A/T278H or cross-linking; however, these antibodies only showed noteworthy binding to the two SOSIP constructs.

Abbreviations

6HB	Six-Helix-Bundle
AA	Amino Acids
Ad5	Adenovirus 5
ADCC	Antibody-Dependent Cellular Cytotoxicity
ADCP	Antibody-Dependent Cellular Phagocytosis
AUC	Area under the Curve
bnAb	Broadly Neutralizing Antibody
BSA	Bovine Serum Albumine
C1-C5	Constant Regions of HIV-1 Envelope
CcdB	Control of Cell Death Toxin B
CCR5	C-C Chemokine Receptor Type 5
CD4	Cluster of Differentiation 4 (receptor)
CD4i	CD4-induced conformation of HIV-1 Env
CDR H3	Complementarity Determining Region 3 of the Antibody Heavy Chain
CRF	Circulating Recombinant Form
CSAM	Cell Surface Antibody Mapping
CT	Cytoplasmic Tail of HIV-1 Envelope
CV	Column Volume
CXCR4	C-X-C Chemokine Receptor Type 4
DMEM	Dulbecco`s Modified Eagle Medium
Dox	Doxycycline
DSC	Differential Scanning Calorimetry
dsDNA	double-stranded Deoxyribonucleic Acid
DSF	Differential Scanning Fluorimetry
E. coli	Escherichia Coli
EC	Extracellular Space
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid
Env	HIV-1 Envelope protein
eOD	Engineered Outer Domain
ER	Endoplasmatic Reticulum
FACS	Fluorescence-Activated Cell Sorting
FC	Fold Change
FCS	Fetal Calf Serum
FDR	False Discovery Rate (corrects p-values for multiple testing)
FP	Fusion Peptide
FPPR	Fusion Peptide Proximal Region
Gag	HIV-1 Group-Specific Antigens (genetic material encoding HIV structural proteins)
gDNA	genomic Deoxyribonucleic Acid
GFP	Green Fluorescent Protein
Gl	Germline
GLA	Glutaraldehyde
gp	Glycoprotein
GT X	Envelope Variant with X Germline-Targeting Mutations
HA	Influenza Hemagglutinin
HEK	Human Embryonic Kidney
HIV-1	Human Immunodeficiency Virus type 1

HR1/2	Heptad Repeats / Helical Regions 1 and 2 of HIV-1 Envelope
HRP	Horse-Raddish Peroxidase
IC	Intracellular Space
IgG	Immunoglobulin G
LB	Lysogeny Broth
mAb	Monoclonal Antibody
MES	2-(N-morpholino)ethanesulfonic acid
MFI	Mean Fluorescence Intensity
MOI	Multiplicity of Infection
MPER	Membrane-Proximal External Region of HIV-1 Envelope
mRNA	messenger Ribonucleic Acid
nAb	Neutralizing Antibody
NFL	Native Flexibly Linked
NGS	Next Generation Sequencing
NHS	N-Hydroxysuccinimide
NIH	National Institute of Health
OD	Optical Density
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PE	Phycoerythrin
PEI	Polyethylenimine
Pen/Strep	Penicillin/Streptomycin
pI	Isoelectric Point
PIC	Pre-Integration Complex
PM	Plasma Membrane
pol	HIV-1 Enzymes (genetic material encoding integrase, protease and reverse transcriptase)
pro	Protease open reading frame
QL	Quick Ligation
Rev	Regulator of Expression of Virion Proteins
sCD4	soluble CD4 (terminated before before the transmembrane domain)
SDS	Sodium Dodecyl Sulfate
SEC	Size Exclusion Chromatography
SHIV	Simian-Human Immunodeficiency Virus (chimeric)
SHM	Somatic Hypermutation
SOSIP	SOS: Disulfide Bridge between gp120 and gp41, IP: I559P Mutation in gp41
SPR	Surface Plasmon Resonance
Tat	Transactivator of Transcription
TB	Terrific Broth
TD	Trimer Derived (HIV-1 Envelope trimer stabilizing mutations
TM	Transmembrane Domain of HIV-1 Envelope
T_M	Melting Temperature
UNAIDS	The Joint United Nations Program on HIV/AIDS
v/v	volume per volume
V1-V5	Variable Loops of HIV-1 Envelope
w/v	weight per volume
WHO	World Health Organization
WT	Wildtype

Danksagung

Hiermit möche ich meine Dankbarkeit all denjenigen gegenüber ausdrücken, die mich und diese Arbeit unterstützt haben.

Zuallererst möchte ich mich bei meinem Doktorvater Prof. Dr. Ralf Wagner dafür bedanken, dass ich meine Doktorarbeit in seiner Arbeitsgruppe durchführen durfte, für all die interessanten Aufgaben die ich bearbeiten konnte und für seine Betreuung, Unterstützung und Anleitung während der letzten vier Jahre. Außerdem möchte ich Prof. Dr. Gernot Längst und Prof. Dr. Ulrike Protzer danken, die als Mentoren im RIGeL Graduiertenprogramm tätig waren und mir mit kritischem Feedback und Ratschlägen zur Seite standen. Ich danke auch Prof. Dr. Dr. André Gessner für seine Unterstützung am Institut für medizinische Mikrobiologie und Hygiene in Regensburg.

Dr. Alexander Kliche und Dr. David Peterhoff bin ich sehr dankbar für ihre Anleitung meiner Labortätigkeit, für die Beantwortung aller Fragen die ich mir während meiner Doktorarbeit gestellt habe und für ihre tolle Unterstützung bei allen Problemen die währenddessen aufgetreten sind. Dr. David Peterhoff hat außerdem alle DSC Messungen die in dieser Arbeit beschrieben sind durchgeführt, wofür ich ihm auch ausgesprochen dankbar bin. Ein besonderes Dankeschön geht auch an Dr. Benedikt Asbach, für all seine Hilfe, Ratschläge und das Programmieren des V3 Analyzers zur Auswertung der FACS Panning Experimente. Auch all meinen anderen Kollegen möchte ich danke sagen, dafür dass sie mich und diese Arbeit immer unterstützt haben, zu den Ergebnissen beigetragen haben und für die stets freundschaftliche und lustige Arbeitsatmosphäre im Labor. Besonders möchte ich dabei allen Studenten danken, die ich während meiner Doktorarbeit betreuen durfte, vor allem Krystina Beer, Matthias Glögl, Christina Schmalzl und Helene Hierl, für ihren unbändigen Fleiß während ihrer Masterarbeiten.

Außerdem möchte ich mich bei Dr. Petra Hoffmann und ihrer Arbeitsgruppe bedanken, vor allem bei Jaqueline Igl, für die Unterstützung am FACS Sorter.

Zuletzt möchte ich mich von ganzem Herzen bei meiner Familie bedanken, vor allem bei meinen Eltern Maria und Siegfried Schmid, die meine Ausbildung sowohl finanziell als auch moralisch immer unterstützt haben, und bei meinem Mann Felix, für seine unendliche Geduld und Hilfe bei allen statistischen Analysen der letzten vier Jahre.

References

1. Barré-Sinoussi, F. *et al.* Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220,** 868–871 (1983).

2. Gallo, R. C. *et al.* Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* **224,** 500–3 (1984).

3. Sharp, P. M. & Hahn, B. H. Origins of HIV and the AIDS epidemic. *Cold Spring Harb. Perspect. Med.* **5,** 1–23 (2011).

4. Hemelaar, J. The origin and diversity of the HIV-1 pandemic. *Trends Mol. Med.* **18,** 1–11 (2012).

5. Buonaguro, L., Tornesello, M. L. & Buonaguro, F. M. Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications. *J. Virol.* **81,** 10209–10219 (2007).

6. WHO Global Health Observatory (GHO) data HIV/AIDS. (2016). Available at: http://www.who.int/gho/hiv/en/. (Accessed: 18th December 2016)

7. UNAIDS Fact Sheet November 2016. (2016). Available at: http://www.unaids.org/en/resources/fact-sheet. (Accessed: 18th December 2016)

8. Archin, N. M., Sung, J. M., Garrido, C., Soriano-Sarabia, N. & Margolis, D. M. Eradicating HIV-1 infection: seeking to clear a persistent pathogen. *Nat. Rev. Microbiol.* **12,** 750–64 (2014).

9. Allers, K. *et al.* Evidence for the cure of HIV infection by CCR5 DELTA 32 / DELTA 32 stem cell transplantation. *Blood* **117,** 2791–2799 (2011).

10. Chun, T.-W. *et al.* Early establishment of a pool of latently infected, resting CD4+ T cells during primary HIV-1 infection. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 8869–8873 (1998).

11. UNAIDS. AIDS by the numbers 2015. Available at: http://www.unaids.org/sites/default/files/media_asset/AIDS_by_the_numbers_2015_en.pdf. (Accessed: 18th December 2016)

12. Bhatti, A. B., Usman, M. & Kandi, V. Current Scenario of HIV/AIDS, Treatment Options, and Major Challenges with Compliance to Antiretroviral Therapy. *Curēus* **8,** e515 (2016).

13. AIDSinfo. Available at: https://aidsinfo.nih.gov/education-materials/fact-sheets/21/55/following-an-hiv-regimen---steps-to-take-before-and-after-starting-hiv-medicines. (Accessed: 1st February 2017)

14. Modrow, S., Falke, D., Truyen, U. & Schätzl, H. *Molekulare Virologie*. (Spektrum, Akad. Verl., 2010).

15. Scherer, L., Rossi, J. J. & Weinberg, M. S. Progress and prospects: RNA-based therapies for treatment of HIV infection. *Gene Ther.* **14,** 1057–64 (2007).

16. Karlsson Hedestam, G. B. *et al.* The challenges of eliciting neutralizing antibodies to HIV-1 and to influenza virus. *Nat. Rev. Microbiol.* **6,** 143–155 (2008).

17. Pritchard, L. K. *et al.* Structural Constraints Determine the Glycosylation of HIV-1 Envelope Trimers. *Cell Rep.* **11,** 1604–1613 (2015).

18. Ward, A. B. & Wilson, I. a. Insights into the trimeric HIV-1 envelope glycoprotein structure. *Trends Biochem. Sci.* **40,** 101–107 (2015).

19. Julien, J.-P., Lee, P. S. & Wilson, I. a. Structural insights into key sites of vulnerability on HIV-1 Env and influenza HA. *Immunol. Rev.* **250,** 180–98 (2012).

20. Kwong, P. D. *et al.* Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393,** 648–659 (1998).

21. Wilen, C. B., Tilton, J. C. & Doms, R. W. HIV: Cell binding and entry. *Cold Spring Harb. Perspect. Med.* **2,** 1–14 (2012).

22. Pancera, M. *et al.* Structure and immune recognition of trimeric pre-fusion HIV-1 Env. *Nature* **514,** 455–61 (2014).

23. Lobritz, M. A., Ratcliff, A. N. & Arts, E. J. HIV-1 entry, inhibitors, and resistance. *Viruses* **2,** 1069–1105 (2010).

24. Wood, N. *et al.* HIV evolution in early infection: Selection pressures, patterns of insertion and deletion, and the impact of APOBEC. *PLoS Pathog.* **5,** (2009).

25. Canducci, F. *et al.* Dynamic features of the selective pressure on the human immunodeficiency virus type 1 (HIV-1) gp120 CD4-binding site in a group of long term non progressor (LTNP) subjects. *Retrovirology* **6,** 4 (2009).

26. Rong, R. *et al.* Role of V1V2 and other human immunodeficiency virus type 1 envelope domains in resistance to autologous neutralization during clade C infection. *J. Virol.* **81,** 1350–1359 (2007).

27. Sagar, M., Wu, X., Lee, S. & Overbaugh, J. Human immunodeficiency virus type 1 V1-V2 envelope loop sequences expand and add glycosylation sites over the course of infection, and these modifications affect antibody neutralization sensitivity. *J. Virol.* **80,** 9586–9598 (2006).

28. Curlin, M. E. *et al.* HIV-1 envelope subregion length variation during disease progression. *PLoS Pathog.* **6,** e1001228 (2010).

29. Mccaffrey, R. A., Saunders, C., Hensel, M. & Stamatatos, L. N-Linked Glycosylation of the V3 Loop and the Immunologically Silent Face of gp120 Protects Human Immunodeficiency Virus Type 1 SF162 from Neutralization by Anti-gp120 and Anti-gp41 Antibodies. **78,** 3279–3295 (2004).

30.   Kwong, P. D. *et al.* HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites. *Nature* **420,** 678–682 (2002).

31.   Lewis, G., Finzi, A., DeVico, A. & Pazgier, M. Conformational Masking and Receptor-Dependent Unmasking of Highly Conserved Env Epitopes Recognized by Non-Neutralizing Antibodies That Mediate Potent ADCC against HIV-1. *Viruses* **7,** 5115–5132 (2015).

32.   West, A. P. *et al.* Structural insights on the role of antibodies in HIV-1 vaccine and therapy. *Cell* **156,** 633–648 (2014).

33.   Overbaugh, J. *et al.* The Antibody Response against HIV-1 The Antibody Response against HIV-1. *Cold Spring Harb. Perspect. Med.* **1,** 1–17 (2012).

34.   Burton, D. R. & Mascola, J. R. Antibody responses to envelope glycoproteins in HIV-1 infection. *Nat. Immunol.* **16,** 571–576 (2015).

35.   Keele, B. F. *et al.* Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 7552–7557 (2008).

36.   Burton, D. R. Antibodies, viruses and vaccines. *Nat Rev Immunol* **2,** 706–713 (2002).

37.   Hraber, P. *et al.* Prevalence of broadly neutralizing antibody responses during chronic HIV-1 infection. *AIDS* **28,** 163–9 (2014).

38.   Moore, P. L., Williamson, C. & Morris, L. Virological features associated with the development of broadly neutralizing antibodies to HIV-1. *Trends Microbiol.* **23,** 204–11 (2015).

39.   Geiß, Y. & Dietrich, U. Catch Me If You Can – The Race Between HIV and Neutralizing Antibodies. *AIDS Rev.* 107–113 (2015).

40.   Yu, L. & Guan, Y. Immunologic basis for long HCDR3s in broadly neutralizing antibodies against HIV-1. *Front. Immunol.* **5,** 28–32 (2014).

41.   Verkoczy, L. & Diaz, M. Autoreactivity in HIV-1 broadly neutralizing antibodies: implications for their function and induction by vaccination. *Curr. Opin. HIV AIDS* **9,** 224–34 (2014).

42.   Zhou, T. *et al.* Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science* **329,** 811–817 (2010).

43.   Falkowska, E. *et al.* PGV04, an HIV-1 gp120 CD4 Binding Site Antibody, Is Broad and Potent in Neutralization but Does Not Induce Conformational Changes Characteristic of CD4. *J. Virol.* **86,** 4394–4403 (2012).

44.   Corti, D. *et al.* Analysis of memory B cell responses and isolation of novel monoclonal antibodies with neutralizing breadth from HIV-1-infected individuals. *PLoS One* **5,** (2010).

45.   Trkola, a *et al.* Human monoclonal antibody 2G12 defines a distinctive neutralization epitope on the gp120 glycoprotein of human immunodeficiency virus type 1. *J. Virol.* **70,** 1100–8 (1996).

46.   Buchacher, A. *et al.* Generation of Human Monoclonal Antibodies against HIV-1 Proteins; Electrofusion and Epstein-Barr Virus Transformation for Peripheral Blood Lymphocyte Immortalization. *AIDS Res. Hum. Retroviruses* **10,** 359–369 (1994).

47.   Frey, G. *et al.* Distinct conformational states of HIV-1 gp41 are recognized by neutralizing and non-neutralizing antibodies. *Nat. Struct. Mol. Biol.* **17,** 1486–1491 (2010).

48.   Huang, J. *et al.* Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. *Nature* **491,** 406–12 (2012).

49.   Sanders, R. W. *et al.* A Next-Generation Cleaved, Soluble HIV-1 Env Trimer, BG505 SOSIP.664 gp140, Expresses Multiple Epitopes for Broadly Neutralizing but Not Non-Neutralizing Antibodies. *PLoS Pathog.* **9,** (2013).

50.   Feng, Y. *et al.* Thermostability of Well-Ordered HIV Spikes Correlates with the Elicitation of Autologous Tier 2 Neutralizing Antibodies. *PLoS Pathog.* **12,** 1–26 (2016).

51.   Sok, D. *et al.* Recombinant HIV envelope trimer selects for quaternary-dependent antibodies targeting the trimer apex. *Proc. Natl. Acad. Sci.* **111,** 17624–17629 (2014).

52.   Scheid, J. F. *et al.* Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. *Nature* **458,** 636–640 (2009).

53.   Doria-Rose, N. A. *et al.* Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. *Nature* **509,** 55–62 (2014).

54.   Walker, L. M. *et al.* Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science* **326,** 285–289 (2009).

55.   Walker, L. M. *et al.* Broad neutralization coverage of HIV by multiple highly potent antibodies. *Nature* **477,** 466–470 (2011).

56.   Pancera, M. *et al.* Crystal structure of PG16 and chimeric dissection with somatically related PG9: structure-function analysis of two quaternary-specific antibodies that effectively neutralize HIV-1. *J. Virol.* **84,** 8098–8110 (2010).

57.   Davenport, T. M. *et al.* Binding interactions between soluble HIV envelope glycoproteins and quaternary-structure-specific monoclonal antibodies PG9 and PG16. *J. Virol.* **85,** 7095–107 (2011).

58.   Andrabi, R. *et al.* Identification of Common Features in Prototype Broadly Neutralizing Antibodies to HIV Envelope V2 Apex to Facilitate Vaccine Design. *Immunity* **43,** 959–973 (2015).

59.   Blattner, C. *et al.* Structural Delineation of a Quaternary, Cleavage-Dependent Epitope at the gp41-gp120 Interface on Intact HIV-1 Env Trimers. *Immunity* **40,** 669–680 (2014).

60.   Huang, J. *et al.* Broad and potent HIV-1 neutralization by a human antibody that binds the gp41-gp120 interface. *Nature* **515,** 138–42 (2014).

61. Sievers, S. a., Scharf, L., West, A. P. & Bjorkman, P. J. Antibody engineering for increased potency, breadth and half-life. *Curr. Opin. HIV AIDS* **10,** 151–9 (2015).

62. Hessell, A. J. *et al.* Broadly neutralizing human anti-HIV antibody 2G12 is effective in protection against mucosal SHIV challenge even at low serum neutralizing titers. *PLoS Pathog.* **5,** (2009).

63. Watkins, J. D. *et al.* An Anti-HIV-1 V3 Loop Antibody Fully Protects Cross-Clade and Elicits T-Cell Immunity in Macaques Mucosally Challenged with an R5 Clade C SHIV. *PLoS One* **6,** e18207 (2011).

64. Postler, T. S. & Desrosiers, R. C. The Tale of the Long Tail: the Cytoplasmic Domain of HIV-1 gp41. *J. Virol.* **87,** 2–15 (2012).

65. Chen, J. *et al.* Effect of the cytoplasmic domain on antigenic characteristics of HIV-1 envelope glycoprotein. *Science (80-.).* **349,** 191–195 (2015).

66. Wan, Y. *et al.* Comparison of immunogenicity between codon optimized HIV-1 Thailand subtype B gp140 and gp145 vaccines. *Vaccine* **25,** 4949–4959 (2007).

67. Heyndrickx, L. *et al.* Selected HIV-1 Env Trimeric Formulations Act as Potent Immunogens in a Rabbit Vaccination Model. *PLoS One* **8,** (2013).

68. Spearman, P. *et al.* A trimeric, V2-deleted HIV-1 envelope glycoprotein vaccine elicits potent neutralizing antibodies but limited breadth of neutralization in human volunteers. *J. Infect. Dis.* **203,** 1165–1173 (2011).

69. Kovacs, J. M. *et al.* HIV-1 envelope trimer elicits more potent neutralizing antibody responses than monomeric gp120. *Proc. Natl. Acad. Sci.* **109,** 12111–12116 (2012).

70. Gray, G. E., Laher, F., Lazarus, E., Ensoli, B. & Corey, L. Approaches to preventative and therapeutic HIV vaccines. *Curr. Opin. Virol.* **17,** 104–109 (2016).

71. Sheets, R. L., Zhou, T. Q. & Knezevic, I. Review of efficacy trials of HIV-1/AIDS vaccines and regulatory lessons learned. A review from a regulatory perspective. *Biologicals* **44,** 73–89 (2016).

72. O'Connell, R. J., Kim, J. H., Corey, L. & Michael, N. L. Human immunodeficiency virus vaccine trials. *Cold Spring Harb. Perspect. Med.* **2,** (2012).

73. Kim, J. H., Excler, J.-L. & Michael, N. L. Lessons from the RV144 Thai Phase III HIV-1 Vaccine Trial and the Search for Correlates of Protection. *Annu. Rev. Med.* **66,** 423–437 (2015).

74. Pitisuttithum, P. HIV vaccine research in Thailand: lessons learned. *Expert Rev. Vaccines* **7,** 311–317 (2008).

75. Du, S. X. *et al.* A directed molecular evolution approach to improved immunogenicity of the HIV-1 envelope glycoprotein. *PLoS One* **6,** (2011).

76. Ching, L. & Stamatatos, L. Alterations in the immunogenic properties of soluble trimeric human immunodeficiency virus type 1 envelope proteins induced by deletion or heterologous substitutions of the V1 loop. *J. Virol.* **84,** 9932–9946 (2010).

77. Selvarajah S1, Puffer BA, Lee FH, Zhu P, Li Y, Wyatt R, Roux KH, Doms RW, B. D. Focused dampening of antibody response to the immunodominant variable loops by engineered soluble gp140. *AIDS Res. Hum. retrovirusesum retroviruses* **24,** 301–14 (2008).

78. Schiffner, T. *et al.* Immune focusing and enhanced neutralization induced by HIV-1 gp140 chemical cross-linking. *J. Virol.* **87,** 10163–72 (2013).

79. Schiffner, T. *et al.* Chemical Cross-linking Stabilizes Native-Like HIV-1 Envelope Glycoprotein Trimer Antigens. *J. Virol.* **90,** JVI.01942-15- (2015).

80. Yasmeen, A. *et al.* Differential binding of neutralizing and non-neutralizing antibodies to native-like soluble HIV-1 Env trimers, uncleaved Env proteins, and monomeric subunits. *Retrovirology* **11,** 41 (2014).

81. Ringe, R. P. *et al.* Cleavage strongly influences whether soluble HIV-1 envelope glycoprotein trimers adopt a native-like conformation. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 18256–61 (2013).

82. Sharma, S. K. *et al.* Cleavage-Independent HIV-1 Env Trimers Engineered as Soluble Native Spike Mimetics for Vaccine Design. *Cell Rep.* **11,** 539–550 (2015).

83. Sanders, R. W. *et al.* HIV-1 VACCINES. HIV-1 neutralizing antibodies induced by native-like envelope trimers. *Science* **349,** aac4223 (2015).

84. Escolano, A. *et al.* Sequential Immunization Elicits Broadly Neutralizing Anti-HIV-1 Antibodies in Ig Knockin Mice. *Cell* **166,** 1445–1458.e12 (2016).

85. Julien, J.-P. *et al.* Design and structure of two HIV-1 clade C SOSIP.664 trimers that increase the arsenal of native-like Env immunogens. *Proc. Natl. Acad. Sci. U. S. A.* **112,** 1–6 (2015).

86. Pugach, P. *et al.* A native-like SOSIP.664 trimer based on an HIV-1 subtype B env gene. *J. Virol.* **89,** 3380–95 (2015).

87. Guenaga, J. *et al.* Structure-Guided Redesign Increases the Propensity of HIV Env To Generate Highly Stable Soluble Trimers. **90,** 2806–2817 (2016).

88. Klasse, P. J. *et al.* Sequential and Simultaneous Immunization of Rabbits with HIV-1 Envelope Glycoprotein SOSIP.664 Trimers from Clades A, B and C. *PLoS Pathog.* **12,** e1005864 (2016).

89. Moody, M. A. *et al.* Strain-Specific V3 and CD4 Binding Site Autologous HIV-1 Neutralizing Antibodies Select Neutralization-Resistant Viruses. *Cell Host Microbe* **18,** 354–362 (2015).

90. Malherbe, D. C. *et al.* Sequential immunization with a subtype B HIV-1 envelope quasispecies partially mimics the in vivo development of neutralizing antibodies. *J. Virol.* **85,** 5262–74 (2011).

91.   Jardine, J. *et al.* Rational HIV immunogen design to target specific germline B cell receptors. *Science* **340,** 711–6 (2013).

92.   Jardine, J. G. *et al.* HIV-1 VACCINES. Priming a broadly neutralizing antibody response to HIV-1 using a germline-targeting immunogen. *Science* **349,** 156–61 (2015).

93.   Briney, B. *et al.* Tailored Immunogens Direct Affinity Maturation toward HIV Neutralizing Antibodies Article Tailored Immunogens Direct Affinity Maturation toward HIV Neutralizing Antibodies. *Cell* **166,** 1459–1464.e11 (2016).

94.   Green, M. R. & Sambrook, J. *Molecular Cloning: A Laboratory Manual.* (John Inglis, Cold Spring Harbor Laboratory Press, 2012).

95.   Bernard P, Gabant P, Bahassi EM, C. M. Positive-selection vectors using the F plasmid ccdB killer gene. *Gene* **148,** 71–74 (1994).

96.   Bruun, T.-H., Mühlbauer, K., Benen, T., Kliche, A. & Wagner, R. A mammalian cell based FACS-panning platform for the selection of HIV-1 envelopes for vaccine development. *PLoS One* **9,** e109196 (2014).

97.   Boussif, O. *et al.* A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proc. Natl. Acad. Sci. U. S. A.* **92,** 7297–7301 (1995).

98.   Bour, S., Boulerice, F. & Wainberg, M. A. Inhibition of gp160 and CD4 maturation in U937 cells after both defective and productive infections by human immunodeficiency virus type 1. *J. Virol.* **65,** 6387–6396 (1991).

99.   Montefiori, D. C. Standardized Assessments of Neutralizing Antibodies for HIV/AIDS Vaccine Development. Available at: https://www.hiv.lanl.gov/content/nab-reference-strains/html/home.htm. (Accessed: 7th February 2017)

100.   NIH AIDS reagents program. Available at: https://aidsreagent.org/reagentdetail.cfm?t=proteins&id=115. (Accessed: 20th May 2017)

101.   Ng, C. T. *et al.* Passive neutralizing antibody controls SHIV viremia and enhances B cell responses in infant macaques. *Nat. Med.* **16,** 1117–1119 (2010).

102.   Wei, X. *et al.* Antibody neutralization and escape by HIV-1. *Nature* **422,** 307–312 (2003).

103.   Gorny, M. K., Xu, J. Y., Karwowska, S., Buchbinder, A. & Zolla-Pazner, S. Repertoire of neutralizing human monoclonal antibodies specific for the V3 domain of HIV-1 gp120. *J. Immunol.* **150,** 635–43 (1993).

104.   Szymczak, A. L. *et al.* Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector. *Nat. Biotechnol.* **22,** 589–594 (2004).

105.   BUCHACHER, A. *et al.* Generation of Human Monoclonal Antibodies against HIV-1 Proteins; Electrofusion and Epstein-Barr Virus Transformation for Peripheral Blood Lymphocyte Immortalization. *AIDS Res. Hum. Retroviruses* **10,** 359–369 (1994).

106.   Yao, F. *et al.* Tetracycline repressor, tetR, rather than the tetR-mammalian cell transcription factor fusion derivatives, regulates inducible gene expression in mammalian cells. *Hum. Gene Ther.* **9,** 1939–50 (1998).

107.   Gossen, M. & Bujard, H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. U. S. A.* **89,** 5547–51 (1992).

108.   Khare, P. D., Rosales, A. G., Bailey, K. R., Russell, S. J. & Federspiel, M. J. Epitope selection from an uncensored peptide library displayed on avian leukosis virus. *Virology* **315,** 313–321 (2003).

109.   Pancera, M. *et al.* Structure and immune recognition of trimeric pre-fusion HIV-1 Env. *Nature* **514,** 455–461 (2014).

110.   Scharf, L. *et al.* Broadly Neutralizing Antibody 8ANC195 Recognizes Closed and Open States of HIV-1 Env. *Cell* **162,** 1379–1390 (2015).

111.   Pantophlet, R. *et al.* Fine mapping of the interaction of neutralizing and nonneutralizing monoclonal antibodies with the CD4 binding site of human immunodeficiency virus type 1 gp120. *J. Virol.* **77,** 642–658 (2003).

112.   Olshevsky, U. *et al.* Identification of individual human immunodeficiency virus type 1 gp120 amino acids important for CD4 receptor binding. *J. Virol.* **64,** 5701–5707 (1990).

113.   Langedijk, J. P. M., Zekveld, M. J., Ruiter, M., Corti, D. & Back, J. W. Helical peptide arrays for lead identification and interaction site mapping. *Anal. Biochem.* **417,** 149–55 (2011).

114.   Quax, T. E. F., Claassens, N. J., Söll, D. & van der Oost, J. Codon Bias as a Means to Fine-Tune Gene Expression. *Mol. Cell* **59,** 149–161 (2015).

115.   Hammonds, J. *et al.* Gp120 stability on HIV-1 virions and Gag-Env pseudovirions is enhanced by an uncleaved Gag core. *Virology* **314,** 636–649 (2003).

116.   Davis, J. & Stacey, G. *Medicines from Animal Cell Culture.* (John Wiley & Sons, Ltd, 2007). doi:10.1002/9780470723791

117.   McLellan, J. S. *et al.* Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature* **480,** 336–343 (2011).

118.   Tran, T. H. *et al.* A Novel High-Throughput Screening Assay to Identify Inhibitors of HIV-1 gp120 Protein Interaction with DC-SIGN. *J. Antivir. Antiretrovir.* **3,** 49–54 (2011).

119.   Zhang, Chung & Oldenburg. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J. Biomol. Screen.* **4,** 67–73 (1999).

120.   Sui, Y. & Wu, Z. Alternative Statistical Parameter for High-Throughput Screening Assay Quality Assessment. *J. Biomol. Screen.* **12,** 229–234 (2007).

121.   Hsu, S.-T. D. & Bonvin, A. M. J. J. Atomic insight into the CD4 binding-induced conformational changes in HIV-1 gp120. *Proteins* **55,** 582–593 (2004).

122. Do Kwon, Y. *et al.* Crystal structure, conformational fixation and entry-related interactions of mature ligand-free HIV-1 Env. *Nat. Struct. Mol. Biol.* **22,** 522–531 (2015).

123. Cordonnier, A., Montagnier, L. & Emerman, M. Single amino-acid changes in HIV envelope affect viral tropism and receptor binding. *Nature* **340,** 571–574 (1989).

124. Cunningham, B. C. & Wells, J. A. Comparison of a structural and a functional epitope. *J. Mol. Biol.* **234,** 554–563 (1993).

125. Hulme, E. C., Bee, M. S. & Goodwin, J. A. Phenotypic classification of mutants: a tool for understanding ligand binding and activation of muscarinic acetylcholine receptors. *Biochem. Soc. Trans.* **35,** 742–745 (2007).

126. Da, L.-T., Quan, J.-M. & Wu, Y.-D. Understanding of the bridging sheet formation of HIV-1 glycoprotein gp120. *J. Phys. Chem. B* **113,** 14536–14543 (2009).

127. Julien, J.-P. *et al.* Asymmetric recognition of the HIV-1 trimer by broadly neutralizing antibody PG9. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 4351–6 (2013).

128. Finzi, A. *et al.* Topological Layers in the HIV-1 gp120 Inner Domain Regulate gp41 Interaction and CD4-Triggered Conformational Transitions. *Mol. Cell* **37,** 656–667 (2010).

129. Finzi, A. *et al.* Conformational characterization of aberrant disulfide-linked HIV-1 gp120 dimers secreted from overexpressing cells. *J. Virol. Methods* **168,** 155–161 (2010).

130. Hoffenberg, S. *et al.* Identification of an HIV-1 clade A envelope that exhibits broad antigenicity and neutralization sensitivity and elicits antibodies targeting three distinct epitopes. *J. Virol.* **87,** 5372–83 (2013).

131. Shaw, T. I. & Zhang, M. HIV N-linked glycosylation site analyzer and its further usage in anchored alignment. *Nucleic Acids Res.* **41,** 454–458 (2013).

132. Gerber, S. *et al.* Mechanism of bacterial oligosaccharyltransferase: In vitro quantification of sequon binding and catalysis. *J. Biol. Chem.* **288,** 8849–8861 (2013).

133. Guenaga, J. *et al.* Well-ordered trimeric HIV-1 subtype B and C soluble spike mimetics generated by negative selection display native-like properties. *PLoS Pathog.* **11,** e1004570 (2015).

134. Heyndrickx, L. *et al.* International Network for Comparison of HIV Neutralization Assays: The NeutNet Report II. *PLoS One* **7,** e36438 (2012).

135. Mata-Fink, J. *et al.* Rapid conformational epitope mapping of anti-gp120 antibodies with a designed mutant panel displayed on yeast. *J. Mol. Biol.* **425,** 444–456 (2013).

136. Grimm, S. K., Battles, M. B. & Ackerman, M. E. Directed evolution of a yeast-displayed HIV-1 SOSIP gp140 spike protein toward improved expression and affinity for conformational antibodies. *PLoS One* **10,** 1–20 (2015).

137. Georgiev, I. S. *et al.* Single-Chain Soluble BG505.SOSIP gp140 Trimers as Structural and Antigenic Mimics of Mature Closed HIV-1 Env. *J. Virol.* **89,** 5318–29 (2015).

138. Zhou, T. *et al.* Transplanting supersites of HIV-1 vulnerability. *PLoS One* **9,** (2014).

139. Gorman, J., Yang, Y., Druz, A., Baxa, U. & Kwong, P. D. Structure-based Design of Trimeric V1V2 Antigens. *AIDS Res. Hum. Retroviruses* **30,** A1–A7 (2014).

140. Hessell, A. J. *et al.* Broadly Neutralizing Monoclonal Antibodies 2F5 and 4E10 Directed against the Human Immunodeficiency Virus Type 1 gp41 Membrane-Proximal External Region Protect against Mucosal Challenge by Simian-Human Immunodeficiency Virus SHIVBa-L. *J. Virol.* **84,** 1302–1313 (2010).

141. Li, Y. *et al.* Mechanism of Neutralization by the Broadly Neutralizing HIV-1 Monoclonal Antibody VRC01. *J. Virol.* **85,** 8954–8967 (2011).

142. Pancera, M. *et al.* Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1-V2-directed antibody PG16. *Nat. Struct. Mol. Biol.* **20,** 804–13 (2013).

143. West, A. P., Diskin, R., Nussenzweig, M. C. & Bjorkman, P. J. Structural basis for germ-line gene usage of a potent class of antibodies targeting the CD4-binding site of HIV-1 gp120. *Proc. Natl. Acad. Sci. U. S. A.* **109,** E2083-90 (2012).

144. McGuire, A. T. *et al.* Engineering HIV envelope protein to activate germline B cell receptors of broadly neutralizing anti-CD4 binding site antibodies. *J. Exp. Med.* **210,** 655–63 (2013).

145. Hoot, S. *et al.* Recombinant HIV Envelope Proteins Fail to Engage Germline Versions of Anti-CD4bs bNAbs. *PLoS Pathog.* **9,** (2013).

146. Dosenovic, P. *et al.* Immunization for HIV-1 Broadly Neutralizing Antibodies in Human Ig Knockin Mice. *Cell* **161,** 1505–1515 (2015).

147. Antibodies, G. B. N. *et al.* Article HIV Vaccine Design to Target Germline Precursors of Article HIV Vaccine Design to Target Germline Precursors of Glycan-Dependent Broadly Neutralizing Antibodies. *Immunity* **45,** 483–496 (2016).

148. Dreyfus, C. *et al.* Highly conserved protective epitopes on influenza B viruses. *Science* **337,** 1343–8 (2012).

149. de Jong, Y. P. *et al.* Broadly neutralizing antibodies abrogate established hepatitis C virus infection. *Sci. Transl. Med.* **6,** 254ra129 (2014).

150. Meola, A. *et al.* Structural flexibility of a conserved antigenic region in hepatitis C virus glycoprotein E2 recognized by broadly neutralizing antibodies. *J. Virol.* **89,** 2170–81 (2015).

151. Gardner, T. J. *et al.* Functional screening for anti-CMV biologics identifies a broadly neutralizing epitope of an essential envelope protein. *Nat. Commun.* **7,** 13627 (2016).

152. Wu, X. *et al.* Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* **329,**

856–61 (2010).

153.   Mouquet, H. *et al.* Complex-type N-glycan recognition by potent broadly neutralizing HIV antibodies. *Proc. Natl. Acad. Sci. U. S. A.* **109,** E3268-77 (2012).

154.   Julien, J.-P. *et al.* Crystal structure of a soluble cleaved HIV-1 envelope trimer. *Science* **342,** 1477–83 (2013).

155.   Forsman, A. *et al.* Llama antibody fragments with cross-subtype human immunodeficiency virus type 1 (HIV-1)-neutralizing properties and high affinity for HIV-1 gp120. *J. Virol.* **82,** 12069–81 (2008).

156.   Strokappe, N. *et al.* Llama antibody fragments recognizing various epitopes of the CD4bs neutralize a broad range of HIV-1 subtypes A, B and C. *PLoS One* **7,** e33298 (2012).

157.   Mufhandu, H. T. *et al.* UCLA1, a synthetic derivative of a gp120 RNA aptamer, inhibits entry of human immunodeficiency virus type 1 subtype C. *J. Virol.* **86,** 4989–99 (2012).

158.   Dey, A. K., Griffiths, C., Lea, S. M. & James, W. Structural characterization of an anti-gp120 RNA aptamer that neutralizes R5 strains of HIV-1. *RNA* **11,** 873–84 (2005).

159.   Pancera, M. *et al.* Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 1166–1171 (2010).

160.   Balla-Jhagjhoorsingh, S. S. *et al.* The N276 Glycosylation Site Is Required for HIV-1 Neutralization by the CD4 Binding Site Specific HJ16 Monoclonal Antibody. *PLoS One* **8,** 8–13 (2013).

161.   Finzi, A. *et al.* Topological layers in the HIV-1 gp120 inner domain regulate gp41 interaction and CD4-triggered conformational transitions. *Mol. Cell* **37,** 656–67 (2010).

162.   AlSalmi, W. *et al.* A New Approach to Produce HIV-1 Envelope Trimers: BOTH CLEAVAGE AND PROPER GLYCOSYLATION ARE ESSENTIAL TO GENERATE AUTHENTIC TRIMERS. *J. Biol. Chem.* **290,** 19780–95 (2015).

163.   Shrivastava, I. & LaLonde, J. M. Fluctuation dynamics analysis of gp120 envelope protein reveals a topologically based communication network. *Proteins* **78,** 2935–49 (2010).

164.   Forsell, M. N. E. *et al.* B cell recognition of the conserved HIV-1 co-receptor binding site is altered by endogenous primate CD4. *PLoS Pathog.* **4,** e1000171 (2008).

165.   Dunn, C. S. *et al.* Human immunodeficiency virus type 1 infection of human CD4-transgenic rabbits. *J. Gen. Virol.* **76 (Pt 6),** 1327–36 (1995).

166.   Keppler, O. T. *et al.* Progress toward a human CD4/CCR5 transgenic rat model for de novo infection by human immunodeficiency virus type 1. *J. Exp. Med.* **195,** 719–36 (2002).

167.   Browning, J. *et al.* Mice transgenic for human CD4 and CCR5 are susceptible to HIV infection. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 14637–41 (1997).

168.   Hatziioannou, T. & Evans, D. T. Animal models for HIV/AIDS research. *Nat. Rev. Microbiol.* **10,** 852–867 (2012).

169.   Melnick, J. L. Advantages and disadvantages of killed and live poliomyelitis vaccines. *Bull. World Health Organ.* **56,** 21–38 (1978).

170.   Rappuoli, R. Toxin inactivation and antigen stabilization: two different uses of formaldehyde. *Vaccine* **12,** 579–81 (1994).

171.   Manji, R. A., Ekser, B., Menkis, A. H. & Cooper, D. K. C. Bioprosthetic heart valves of the future. *Xenotransplantation* **21,** 1–10 (2014).

172.   Dardik, H. *et al.* Comparative decades of experience with glutaraldehyde-tanned human umbilical cord vein graft for lower limb revascularization: an analysis of 1275 cases. *J. Vasc. Surg.* **35,** 64–71 (2002).

173.   Riechelmann, H., Schmutzhard, J., van der Werf, J. F., Distler, A. & Kleinjans, H. A. J. Efficacy and safety of a glutaraldehyde-modified house dust mite extract in allergic rhinitis. *Am. J. Rhinol. Allergy* **24,** e104-9

174.   García-Sellés, J. *et al.* Clinical efficacy and safety of a depigmented and glutaraldehyde polymerized therapeutic vaccine of Parietaria judaica. *Allergol. Immunopathol. (Madr).* **31,** 63–9

175.   Park, S.-N., Park, J.-C., Kim, H. O., Song, M. J. & Suh, H. Characterization of porous collagen/hyaluronic acid scaffold modified by 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide cross-linking. *Biomaterials* **23,** 1205–12 (2002).

176.   Gomes, R. *et al.* KSAC, a defined Leishmania antigen, plus adjuvant protects against the virulence of L. major transmitted by its natural vector Phlebotomus duboscqi. *PLoS Negl. Trop. Dis.* **6,** (2012).

177.   Coler, R. N. *et al.* A synthetic adjuvant to enhance and expand immune responses to influenza vaccines. *PLoS One* **5,** (2010).

178.   Noe, A. R. *et al.* A full-length Plasmodium falciparum recombinant circumsporozoite protein expressed by Pseudomonas fluorescens platform as a Malaria vaccine candidate. *PLoS One* **9,** (2014).

179.   O'Gorman, S., Fox, D. T. & Wahl, G. M. Recombinase-mediated gene activation and site-specific integration in mammalian cells. *Science* **251,** 1351–1355 (1991).

180.   Beyer, D. Multiplication of HIV. (1997). Available at: https://commons.wikimedia.org/wiki/File:Hiv_gross_german.png. (Accessed: 19th December 2016)

181.   Wang, H. *et al.* Cryo-EM structure of a CD4-bound open HIV-1 envelope trimer reveals structural rearrangements of the gp120 V1V2 loop. *Proc. Natl. Acad. Sci. U. S. A.* **113,** E7151–E7158 (2016).