Exploring HIV-1 - Host Cell Interactions using Peptides derived from Human Pegivirus-1, and from Combinatorial Libraries

Der Naturwissenschaftlichen Fakultät

der Friedrich-Alexander-Universität Erlangen-Nürnberg

zur

Erlangung des Doktorgrades Dr. rer. nat.

vorgelegt von

Rebecca Hoffmann

aus Potsdam

Als Dissertation genehmigt

von der Naturwissenschaftlichen Fakultät

der Friedrich-Alexander-Universität

Erlangen-Nürnberg

Tag der mündlichen Prüfung:	03.12.2018
Vorsitzender des Promotionsorgans:	Prof. Dr. Georg Kreimer
Gutachterinnen:	Prof. Dr. Jutta Eichler
	Prof. Dr. Barbara Schmidt

Meinen Eltern

Content

Abbreviations

AA	amino acid
AcOH	acetic acid
Ahx	6-Aminohexanoic acid
AIDS	acquired immune deficiency syndrome
Alloc	Allyloxycarbonyl
AOA	3,6-dioxaoctanoic acid
aq	aqueous solution
ART	antiretroviral therapy
AVG	average
Bio	Biotin
bnAb	broadly neutralizing antibody
Boc	*tert*-Butoxycarbonyl
C	concentration
c.f.	lat. confer meaning compare
CCR5	C-C chemokine receptor type 5
CD4	cluster of differentiation 4
CD4bs	CD4 binding site
conc.	concentrated
CRF	circulating recombinant form
CXCR4	C-X-C chemokine receptor type 4
Dap	2,3-diaminopropionic acid

DCM	Dichloromethane
DIC	*N,N'*-Diisopropylcarbodiimide
dil.	diluted
DIPEA	*N,N*-Diisoproplyethylamine
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DTNB	5,5'-Disulfanediylbis(2-nitrobenzoic acid)
DTNB	5,5´-Disulfaneddiylbis(2-nitrobenzoic acid) (Ellmann's reagent)
ELISA	enzyme-linked immunosorbent assay
ENV	envelope glycoprotein
eq.	equivalent
ESI	electrospray ionization
FBS	fetal bovine serum
fig.	figures
Fluo	Fluorescein
Fmoc	Fluorenyl-9-methyloxycarbonyl
FOX	Fold over X
gp120	glycoprotein 120 kDa
gp41	glycoprotein 41 kDa
GuHCl	Guanidinium hydrochloride
Hcy	Homocysteine
HIV	human immunodeficiency virus
HPgV-1	human pegivirus type 1

HPLC	high-performance liquid chromatography
HRP	horse radish peroxidase
logP	partition coefficient
m/z	mass ion charge ratio
mAb	monoclonal antibody
MPER	membrane proximal external region
n.d.	not determined
NK cells	natural killer cells
OAll	Allyl
OD	optical density
OD_{492nm}	optical density at 492nm
OPD	
OtBu	*tert*-butyl
ox	oxidized
Oxyma Pure	Ethyl cyano(hydroxyimino)acetate
PB	phosphate buffer pH 7.2
Pbf	2,2,4,6,7-Pentamethyldihydrobenzofuran-5-sulfonyl
PBMC	peripheral blood mononuclear cells
PP	Polypropylene
PP	Polypropylene plate
PS	positional scanning
PS-SCL	positional scanning synthetic combinatorial library
PyBOP	Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate

RT	room temperature
SCLs	synthetic combinatorial libraries
SD	standard deviation
SE	standard error
SS	stock solution
TCEP	Tris(2-carboxyethyl)phosphine
TFE	2,2,2-Trifluoroethanol
TPIMS	Torrey Pines Institute for Molecular Studies
Tris	2-Amino-2-(hydroxymethyl)propane-1,3-diol
Trt	Trityl
w/o	without
X	Amino hexanoic acid

1 Introduction

1.1 Peptides: Versatile Molecular Tools

In the merging fields of pharmaceutical and chemical biology research, peptides are used as lead compounds for drug discovery, protein binding site mimetics, as tools for the investigation of ligand binding properties, or as drug itself [1, 2]. More than half a decade has passed since the breakthrough development of the solid phase peptide synthesis (SPPS) [3]. This method has been optimized continuously, and state of the art SPPS facilitates convenient rapid, parallel and automatable synthesis of high quality products [4].The simultaneous synthesis, the modular combination of amino acids as well as synthetical modified derivates as building blocks and their attractive compatibility with usually applied testing systems and *in vivo* are important advantages for choosing peptides over small molecules [1, 5]. The synthesis approach as well as the amid bond structure combine the ability to create uniquely featured compounds with a great applicability. Attaching additional building blocks like sugars of fatty acids or using the peptide itself as an attachment for labelling another compound are only two examples for the using modified peptides in the fields of molecular and chemical biology [4]. Moreover, naturally occurring peptides play a crucial role in different physiological processes and the usage of endogenous compounds derived from proteins usually exhibits negligible toxicity if applied to humans [5]. The pathogenesis of diseases is often related to receptor mediated processes, or other protein-involving mechanisms. Thus, protein - protein interactions are a main focus for investigation possible targets for prevention or treatment of diseases. Peptides have proven their suitability as protein binding site mimetics in various fields. The versatility of displaying any protein fragment structure as an exact copy or incorporating diverse side chain modification highlights the advantageous properties of peptides as epitope or binding site mimetics [6]. Moreover, peptides can be used to present linear or cyclic binding site, as well as combining several components of a discontinuous binding site [2]. Cyclic binding site mimetics, which can fold in a stable β-hairpin conformation, were reported as potent trypsin inhibitors, or showed high affinity binding to C-X-C-chemokine receptor type 4 (CXCR4) [7]. In the field of HIV-1 research, the highly active α-helical 35mer peptide fragment of the gp41 subunit of the envelope protein, called T-20 or Enfuvirtide, is a FDA approved entry inhibitor[8, 9]. Promising in

vitro results were also shown for a peptide that is a mimicking the paratope of the neutralizing antibody b12 [10]. The mimicry of the discontinuous binding site of CXCR4 using a peptide resulted in a functional selective HIV-1-inhibitor [11-13]. Peptides with their outstanding functional role as binding site mimetics are versatile tools to investigate protein-protein-interactions.

1.2 Human Immunodeficiency Virus 1 (HIV-1)

1.2.1 Global Virus Spread

The *human immunodeficiency virus* is the known pathogen causing the acquired immune deficiency syndrome (AIDS). Despite the decline in public awareness, HIV is still causing an epidemic infection with around 37 million HIV positive individuals in 2016. About 5000 people get infected daily, whereby over 50 % new infection were registered in sub-Saharan Africa. Nevertheless, the total number of new infections has declined by 16 % compared to 2010. This global decline is a great success, but an alarming uprising trend is observed in Eastern Europe and Central Asia over the past five years [14]. Besides that, the restricted access to antiretroviral therapy is still not solved and resulting in approximately half of the infected individuals still not receiving medicinal treatment [15]. The current therapy is usually a lifelong triple drug combination called ART (antiretroviral therapy). The therapeutic approach combines drugs addressing different target sites in the viral replication cycle (entry, transcription, integration and virus assembly) [16]. The vast majority of worldwide infections are caused by HIV-1 but there is a less pathogenic virus, called HIV-2, circulating predominantly in West Africa. Until today, HIV-2 infection is basically restricted to this specific region. Both viruses belong to the subgenus primate lentivirus as does the *simian immunodeficiency virus* (SIV), which infects non-human primates [17]. The generally accepted hypothesis proposes that HIV evolved through SIV cross-species infection [18]. The *primate lentiviruses* are single stranded RNA viruses; a subgenus belonging to the *retroviridae*. The virus particles present their envelope protein (ENV) as spikes on the surfaces and several embedded proteins are enclosed together with the viral RNA by the capsid envelope (Figure 1).

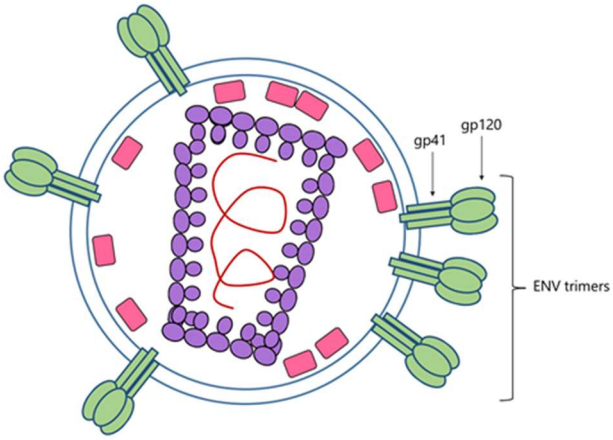


Figure 1: Schematic of HIV-1 particle structure. ENV trimers are presented on the surface. adapted from [19].

The viral replication, using human CD4 (cluster of differentiation 4)-expressing host cells, can be divided in six steps: Entry, Reverse Transcription, Integration, Transcription, Virus assembly and Budding[19]. Entry, Reverse Transcription, Integration and Virus assembly steps are targets of FDA approved drugs. The combination of these drugs allows an effective treatment of the infection, although a lifelong therapeutic regimen is mandatory [16]. So far, no broadly applicable cure or vaccination is available. The tremendous number of people living with HIV-1, as well as new infections per year are demonstrating the importance of continuous research to understand the molecular determinants in detail towards the goal of an effective HIV-1 infection prevention and cure.

1.2.2 HIV-1 ENV mediated Host Cell Entry

HIV-1 replication is initiated by the host cell entry which proceeds via a sophisticated mechanism. The surface- presented viral glycoprotein is of particular interest as it mediates the entry process. The mature HIV-1 envelope protein (ENV) is composed of two subunits, glycoprotein 120 (gp120) and glycoprotein 41 (gp41). ENV units are associated as trimers via noncovalent interaction and anchored into the viral membrane by the transmembrane region of gp41 [20]. The gp120 sequence consists of five conserved (C1-C5) and four variable (V1-V4) domains (Figure 2) [21]. The variable domains are also referred to as loops.

Figure 2: Schematic of gp120 protein sequence. The sequence contains variable (**V1-V5**) and conserved (*C1-C5*) regions, adapted from [21].

The second ENV subunit, gp41, consists of a surface accessible ectodomain and a membrane spanning part. The ectodomain can be further divided into several functional domains: fusion peptide, N-heptad and C-heptad repeat, disulfide loop and membrane proximal external region (MPER). The transmembrane domain and the cytoplasmic tail, belonging to the membrane spanning part, are buried inside the viral particle (Figure 3) [21, 22].

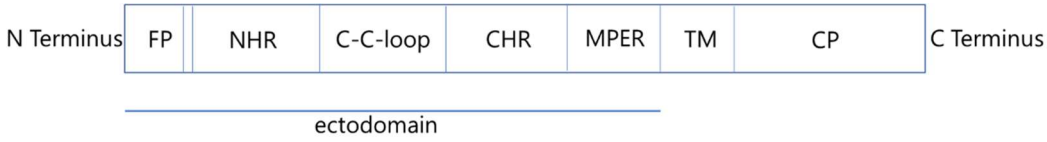


Figure 3: Schematic of gp41 protein sequence. The sequence contains several functional domains: FP = fusion peptide, NHR = N-heptad repeat, C-C-loop = disulfide loop, CHR = C-heptad repeat, MPER = membrane proximal region, TM = transmembrane domain, CP = cytoplasmic tail, adapted from [22].

For efficient host cell entry, a fully matured ENV is mandatory. The HIV-1 viral spike has evolved a complex mechanism for attachment, entry and subsequent fusion with host cell membranes to facilitate the entry. Major determinants and steps have been identified, but the process as a whole is not yet understood in detail. The key principle of the fusion process is a succession of conformational changes of ENV. Prior to the interaction with the host cell receptors, ENV is in the pre-fusion state. HIV-1 virus particles infect predominantly CD4- presenting cells by using CD4 as the its main receptor for cell entry. The initial conformational change is introduced by binding of gp120 to CD4 on the host cell surface. Binding to CD4 causes conformational changes within V1 and V2 domain of gp120. The interaction sites with CD4 are located in distinct region of gp120 called CD4 binding loop and in parts of C4, V5 and C5, forming the discontinuous CD4-binding site [23]. Subsequently, V3 loop rearrangement elicits the binding to a host cell presented receptor. CXCR4 (C-X-C chemokine receptor type 4) and CCR5 (C-C chemokine receptor type 5) are the main coreceptors addressed by HIV-1. HIV-1 strains can be classified by their receptor-dependent tropism as R5 (CCR5), X4 (CXCR4) and R5X4 (dual) tropic [24]. The V3 loop sequence of gp120 was identified as the molecular determinant

for coreceptor selectivity of HIV-1 strains. The V3 loop positive net charge is proposed to be crucial for receptor tropism. Moreover, the V3 domain is an accessible target for HIV-1 neutralizing antibodies and involved in sensitivity decrease against the FDA approved entry inhibitor Enfuvirtide. [25, 26]. The V3 loop, a short section of gp120 with a length of usually 35 AA, is a key element for coreceptor selectivity and binding and consequently outstandingly important for the entry process. After coreceptor binding a cascade of conformational changes starts. As a result, the fusion peptide of gp41 is inserted into the host cell membrane and followed by gp41 hinging the disulfide loop under six-helix-bundle formation. The two heptad repeats of gp41 are forming a bundle which enables the fusion pore formation due to close proximity of both membranes. The viral RNA and proteins are transferred to host cell cytoplasm *via* the formed pore [24]. The entry steps are depicted in Figure 4.

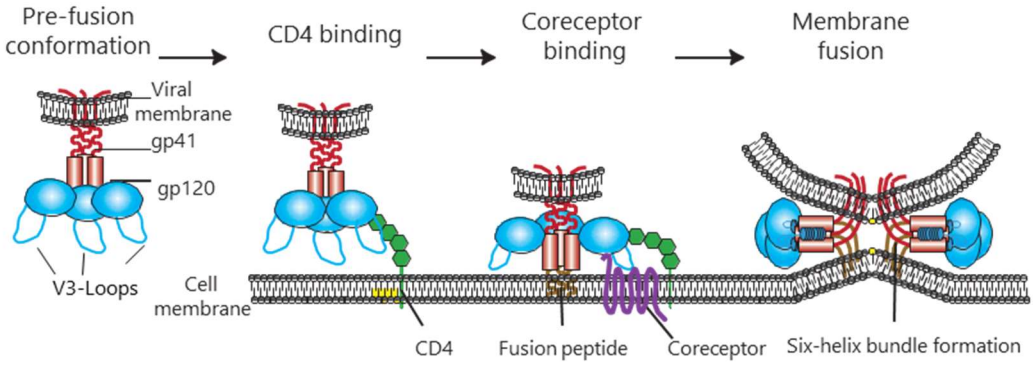


Figure 4: Schematic of the HIV-1 entry process. HIV-1 entry includes several steps and conformational changes: CD4 binding to gp120, coreceptor binding to V3 Loop of gp120 and membrane fusion after six helix bundle formation of gp41 adapted from [24].

1.2.3 Broadly Neutralizing HIV-1 Antibodies

Since the early beginnings of research in the fields of HIV-1 infection and AIDS treatment, a major focus was the investigation of antibodies targeting the viral spike. The humoral response during HIV-1 infection was intensively studied [27]. Antibodies isolated from sera of HIV-1 infected individuals showed neutralizing activity with efficacy levels ranging from low to broadly neutralizing antibodies. A key interest was to identify antibodies with cross-clade neutralization potency and subsequently investigate the suitability for passive immunization

strategies [28, 29]. The second main approach was focused on the generation of broadly neutralizing antibodies (bnAbs) by using an effective immunogen. Vaccine strategies using a potent immunogen were extensively investigated up to the stage of clinical trials with gp120 constructs e.g. the *Vax004 trial* [30-34].

The importance of bnAbs for HIV-1 research is not limited to the field of vaccine development. Moreover, the study of bnAbs provided new insight into the HIV-1 entry mechanism, especially stable ENV conformations, binding modes and specific ENV vulnerability sites [35]. The intensified search for broadly neutralizing antibodies revealed only a few specific binding epitopes which were highly conserved across clades and accessible for humoral response with the latter being the most severe impairment for this approach. The conserved and therefore "vulnerable" core structures of viral envelope protein are highly shielded [36, 37]. For example, the extensive glycosylation covering the ENV surface is leaving only 3 % theoretically accessible for antibodies [38]. So far, only a few target sites for broadly neutralizing antibodies are described (Figure 5). The MPER region of gp41 is addressed by antibodies 2F5, Z13e1, 4E10 and 10E8, which are distinctly targeting the six-helix-bundle formation during the entry process. The antibody epitopes are all located in the MPER (AA 662-684 HXBc2 numbering) and partly overlapping. Despite the shared target epitope each antibody has a unique mechanism for neutralization [39, 40]. For instances monoclonal antibody (mAb) 4E10 interacts additionally with lipid bilayers of the membrane, which is assumed to be crucial for its broad neutralizing potency [41, 42]. The glycosylation of ENV, which is shielding the protein surface, is the described epitope for mAbs isolated from HIV-1 infected individuals. One of the most potent antibodies targeting glycans is called 2G12 with an specific glycan epitope in V3 loop region of gp120 [43]. A conformation of the V1/V2 site of gp120 was further discovered as a target of a new class of bnAbs. Two representatives are PG9 and PG16, which are classified as broad neutralizers [44-46]. The CD4 binding site (CD4bs) of gp120 overlaps with the epitopes of a group of neutralizing antibodies. The discovery of the unique binding mode of mAB b12 was a milestone in the HIV-1 entry inhibition research. The discontinuous epitope partly overlaps with the CD4 binding site [47]. The epitope of another even more potent bnAb, called VRC01, is also located within the CD4 binding site [48, 49]. The isolation of VRC01 was the starting point for characterization of a whole class of neutralizing antibodies including VRC03 and 3BNC117 [50, 51]. A phase 1 clinical trial with VRC01 has already been completed and the preparation for a

follow up is currently ongoing [52, 53]. Recently a new CD4bs antibody was published with near pan neutralization potency [54]. This outstanding potency illustrates the potential of the CD4 binding site of gp120 as target for entry inhibition.

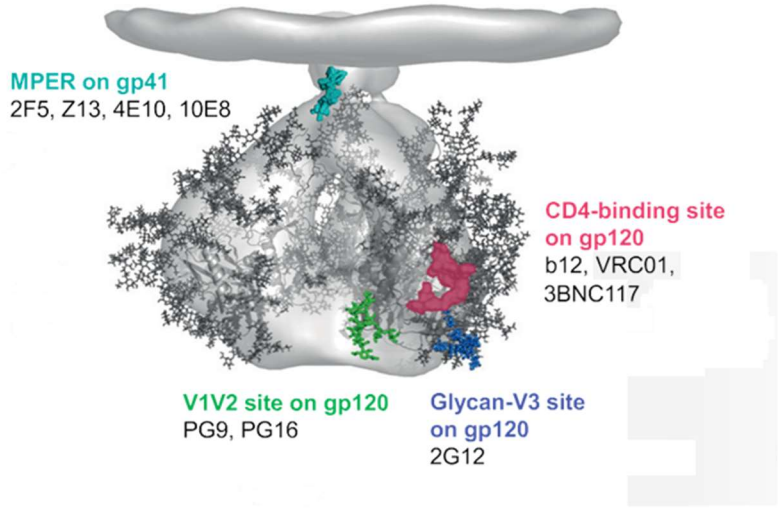


Figure 5: HIV-1 viral spike model with glycosylation and highlighted sites of targeted ENV binding site of listed broadly neutralizing antibodies. red = CD4 binding site on gp120, blue = Glycan-V3 site on gp120, green = V1/V2 site on gp120 and cyan = MPER on gp41, adapted from [29].

The characterized neutralizing antibodies mainly target defined epitopes of ENV. Moreover, potent neutralizing antibodies with additional epitope features were recently reported like a trimer specific antibody targeting a gp120-gp41 interface [55]. A trispecific antibody mediated effective infection protection in macaques. This antibody combines MPER of gp41, V1-V2 region and the CD4 binding site of gp120 as a discontinuous epitope [56]. The broadly neutralizing antibodies illustrate that the viral spike has only a small set of vulnerable sites. Each site has its unique features and the accessibility is highly dependent on ENV conformation.

1.2.4 Nomenclature: Groups and Group M Subtypes

The worldwide HIV-1 pandemic is not caused by a single virus strain. It is a rather great diversity of strains with different spreads and varying degrees of relationship. Therefore, a nomenclature system was established for categorization of the strains circulating worldwide. The system is based on phylogenetic cluster analysis. The different HIV-1 lineages are sorted in three groups: M, N and O. Group M is the major cause for worldwide infection and can be further divided

into clades which are called subtypes. The classification to a certain subtype (e.g. A, B, C,) depends on the genetic distance. Likewise, sub-subtypes (e.g. A1, A2) can also be classified. Currently, 10 subtypes are defined and numbered from clade A to clade K. The genetic variability between subtypes of group M for envelope protein (ENV) genes is approximately 25 %. In addition, the subtype classes of Circulating Recombinant Form (CRF) are defined. One recombinant lineage shares the same mosaic structure. For example, CRF-AG shares the same set of genes belonging to clade A and clade G [57, 58]: Figure 6 illustrates the relationship of the three HIV-1 groups and subtypes and CRFs of group M.

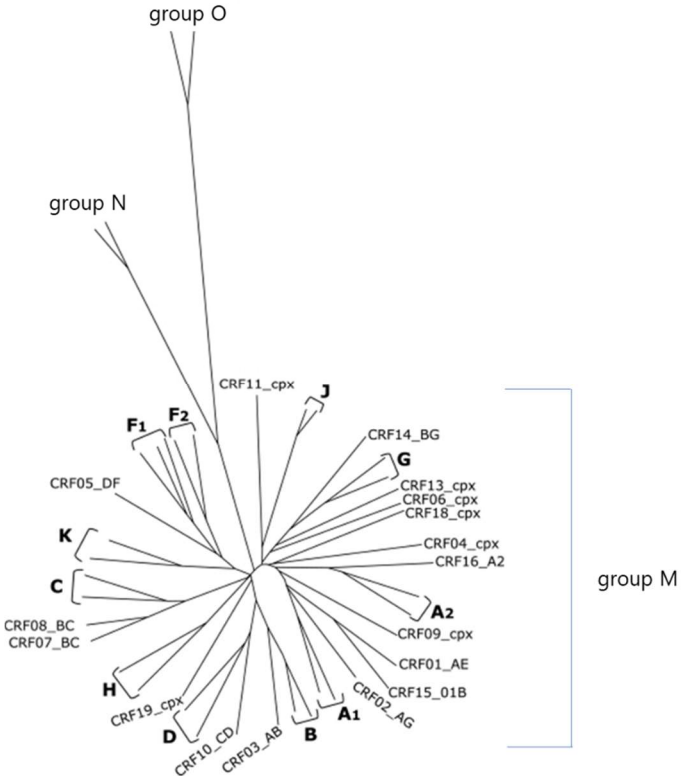


Figure 6: Phylogenetic tree showing HIV-1 groups M, N, O. Group M is divided into subtypes and CRFs. The phylogenetic analysis is based on the neighbor-joining method and adapted from [58].

1.3 The Human Pegivirus Type 1

1.3.1 A Non-pathogenic Worldwide Distributed Virus

The human Pegivirus type 1 (HPgV-1) was formerly known as GBV-C or hepatitis G virus. The bloodborne virus was discovered in human serum samples from two different laboratories in the 1990s [59-61]. The RNA virus belongs to a newly defined genus *Pegivirus* of the *Flaviviridae* family[62]. Although it was initially named hepatitis virus G, HPgV-1 is not causing any viral hepatitis infection in humans [60, 63]. Moreover, no human disease is yet related to the HPgV-1 infection, which until now is considered a non-pathogenic human virus. The double stranded RNA virus infects both immune-competent and deficient humans. It is estimated that 25 % of infections persist longer than two years. Prevalence of viraemia varied between 1-5 % in industrial countries and 20 % in developing countries, measured in samples from healthy blood donors. Studies are estimating that around 750 million people are living with acute HPgV-1 infection. Other studies which are using anti-viral antibodies as a marker for infection suggest that approximately 1.5-2.5 billion infected people, either acute or post infected, are currently living worldwide. The virus is transmitted sexually, vertically and by exposure to infected blood products. Very little is known about its host cell entry mechanism and general replication cycle due to the limited in vitro data. Early studies were suggesting a hepatotropic pathway, yet latest studies are proposing a lymphotropic one. HPgV-1 RNA was detected in spleen, bone marrow, liver, cerebrospinal fluid and peripheral blood mononuclear cells (PBMCs). There is evidence that indicates that PBMCs are the plausible replication site [64-66]. HPgV-1 has a single stranded RNA positive sense genome, encoding several nonstructural (NS) and two envelope proteins (E) (Figure 7). These two envelope glycoproteins, called E1 and E2, are assumed to be presented as heterodimers on the viral spike surface [64, 65].

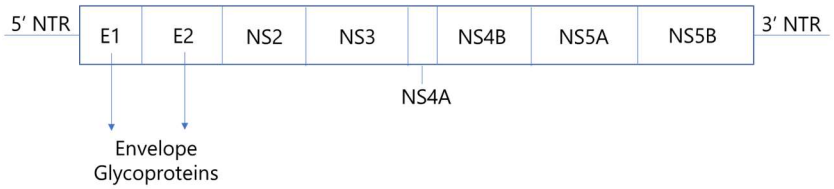


Figure 7: Schematic of the HPgV-1 genome: Single-stranded RNA genome with a 5′ and 3′ non-translated region (NTR). NS = nonstructural protein E = envelope protein, adapted from [64].

The worldwide circulating HPgV-1 isolates can be divided in at least five subtypes with regional distribution patterns. Subtype 2 is the predominant subtype in Europa and America [67-70]. Additionally, three different genotypes were proposed, as well as further subdivision of subtype 1 and 2 (Table 1) [71-73].

Table 1: HPgV-1 subtype and regional prevalence

subtype	region
1	West Africa
2	America, Europe
3	Asia
4	Southeast Asia
5	South Africa
6	Indonesia
7*	Yunman Province China
8**	Pacific

*proposed by [71]

** proposed by [73]

HPgV-1 infections could be detected in the human population all over the world, including indigenous tribes in Papua New Guinea, Central and South America with regional subtype prevalence, and the hypothesis can be concluded that the virus evolved together with its

human host [65, 74]. It can be assumed that the HPgV-1 virus is an ancient companion of humankind with no directly pathogenic impact caused by its infection.

1.3.2 Viral Interference with HIV-1

The transmission routes of HPgV-1 are overlapping with other human viruses like HIV-1. As a consequence, studies confirmed a higher prevalence of HPgV-1 in HIV-1 positive individuals. Depending on the study, HPgV-1 viraemia varies between 16-39 %. Moreover, the predominant number of epidemiological studies showed immunological evidence that a co-infection was beneficial for HIV-1 positive individuals. Different outcome parameters like disease progression, mortality or CD4-cell counts were analyzed and demonstrated a significant difference between HPgV-1 infected and non-infected individuals [65, 75]. This phenomenon is called viral interference however does not define the underlying mechanism that is causing the effect [76]. Based on the same transmission routes and shared host cells, it can be expected that multiple pathways are responsible for the beneficial effects of a HPgV-1 infection in HIV-1 disease progression [65]. The inhibition of HIV-1 strains in vitro strongly indicates a direct inhibitory effect of HPgV-1 on HIV-1 replication [77]. Over the past twenty years, several molecular mechanisms were postulated that could be causing viral interference with HIV-1 [75]. Laboratories demonstrated an involvement of HPgV-1 proteins in HIV-1 inhibition. Moreover, a distinct domain of the nonstructural protein NS5A and the glycoprotein E2 of HPgV-1 were described to inhibit the HIV-1 entry to its host cell in vitro [78, 79]. The envelope protein E2 is predicted to be involved in HPgV-1 host cell entry *via* a receptor-based mechanism. So far, the binding of recombinant E2 protein to human and murine cell lines could be shown [80, 81]. The used receptor remains unidentified. The HPgV-1 RNA was detected in T and B lymphocytes, natural killer (NK) cells and monocytes of infected patient sera [82]. In addition, a study with HIV-1 infected individuals showed a HPgV-1 infection depending effect on HIV-1 coreceptor expression on CD4 positive cells. Coinfected individuals had a reduced CXCR4 surface expression but CCR5 up regulation in advanced stages of the infection was not observed. This might be indicating that HPgV-1 is interfering with the HIV-1 host cell coreceptors CXCR4 and CCR5 [83].

1.3.3 E2 Protein derived Peptides Inhibit HIV-1 Infection in Vitro

The envelope glycoprotein 2 (E2) of HPgV-1 interferes in vitro with early steps of the HIV-1 replication cycle [79]. Previous work demonstrated the inhibitory activity of a recombinant variant of the E2 protein against HIV-1 isolates in vitro [79]. The recombinant E2 protein presented the proposed cytoplasmic domain (AA 1-340) and was equally effective as the full length E2 protein [84].The mode of action and inhibitory region of E2 remained unclear. That is why the E2 sequence was screened using overlapping peptide sequences representing the envelope protein of HPgV-1. The screening identified inhibitory peptides against X4- and R5-tropic virus isolates with IC_{50} values in the lower micromolar range. The inhibitory peptide sequences were located within the N-terminal part of the E2 sequence. The selectivity was tested with a selected panel of peptides. These peptides were not active against SIV and HIV-2 in neutralization assays with TZM-bl cells [85]. Two representatives of active peptides inhibited the virus-cell fusion in a CD4 independent manner. In PBMCs and TZM-bl based assays no remarkable effect on CD4 or coreceptor expression was observed. Two active lead peptides with a 6mer sequence overlap were characterized as most potent (Figure 8) [22].

$_{37}$WDRGNVTL**LCDCPNGP**WVWVPAFCQAVG$_{64}$

Figure 8: E2 protein sequence part representing the active lead peptides identified by Reil laboratory [22, 85]. bold = overlapping sequence of the two active peptides. Iowa isolate sequence and numbering (Genbank Accession Number: AF121950).

Beside the inhibitory N-terminal peptides of E2, another peptide sequence was identified as inhibitory against HIV-1 replication in vitro and NK cell function [86, 87]. Peptides derived from the envelope 1 glycoprotein were also reported as potent in cell fusion assays [88]. The E2 sequence contains several inhibitory peptides which were active in a number of in vitro HIV-1 neutralization assays with laboratory and clinical HIV-1 isolates.

1.4 Mixture based Peptide Combinatorial Libraries

Medicinal basic research utilizes several techniques for *de novo* identification of compounds that interfere with known pathogenesis pathways e.g. based on protein-protein-interactions. Due to the large diversity of synthetic and physiological occurring compounds, single compound testing is limited to reasonable time and budget. To overcome this limitation combinatorial concepts were developed. Each one is based on an unique technique to generate compound libraries [89]. Exemplary a synthesis approach was developed called teabag method which in particular facilitates rapid parallel peptide synthesis [90]. Using this approach large peptide libraries up to billions of individual compounds can be synthesized. The base of this method is to screen compound mixtures to subsequently identify active individual compounds [91]. These mixtures are named synthetic combinatorial libraries (SCLs). Combinatorial peptide libraries ranging from three to ten amino acids in length have been successfully synthesized and applied in several in vitro screening assays. Key benefits of peptide SCLs are the solubility, the compositional diversity and the possibility of a label-free application. Mixture-based libraries can be arranged in a positional scanning format (PS-SCL) and the incorporation of L- and D- as well as unnatural amino acids is facilitated by a solid phase synthesis approach[92]. A PS-SCL is arranged in sublibraries. The number of sublibraries depends on the length of the sequences. Each sublibrary represents the same set of peptides, but in a different arrangement. For further illustration, a pentapeptide (N-terminal acetylated, C-terminal amide) PS-SCL will be exemplary elucidated. In theory for this example 20 different amino acids are incorporated per position. A pentapeptide library is composed of five sublibraries with 20 mixtures per sublibrary. Each sublibrary has one amino acid position defined (O). Instead of parallel coupling 20 different AAs at this defined position, only one AA is specifically incorporated (O). Consequently, the sublibraries vary in the position of the defined position. The arrangement of the mixture is the only difference because each sublibrary contains the same set of peptides. The whole PS-SCL is composed of 100 mixtures. These 100 PS-SCL are tested in the dedicated assay of interest. The results provide information about the most active amino acid per position. Additional information regarding the activity of each functionality can be gained by the analysis of the PS-SCL screening results [93, 94]. The testing of the 100 PS-SCL mixtures enables the deconvolution of individual peptides by firstly selecting the most active AA per position and secondly combining the selection to generate all combinations for possible

individual peptides (Figure 9). To determine the activity of the individual peptides, they are synthesized and then tested in the dedicated assay of interest.

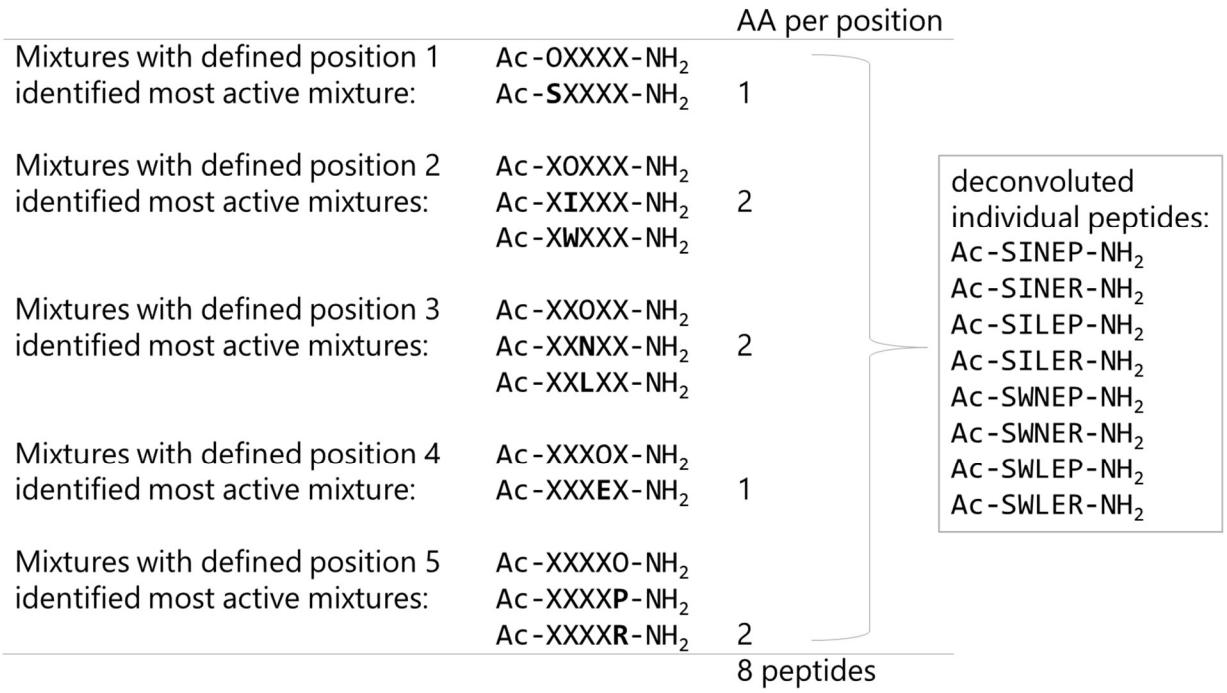


Figure 9: Scheme of a positional scanning process of a pentapeptide PS-SCL library and deconvolution of the most active eight peptides, adapted from [94].

Using PS-SCL is an established approach for screening a large number of compounds. It is especially useful when the number of samples testable is limited due the design of the assay of interest. The screening of PS-SCLs resulted in active compounds which for example targeted opioid receptors or enzymes [95]. Prior to PS-SCL screening with the designated assay, an initial verification of the library activity is mandatory. A strategy to identify active libraries is to rank by scaffold activity. Scaffold ranking samples (also referred to as *all X samples*) of a PS-SCL have no position defined and is a mixture of all positional scanning mixtures. This allows two rank SCLs based on their scaffold activity. Consequently, a panel of scaffold samples can be tested with the assay of interest. The ranking by scaffold activity will be further used to select libraries and subsequently to be applied in-depth screenings. The scaffold ranking approach has proven its utility for rapid identification of active libraries [95].

2 Aim of work

The aim of this work was to use peptides for a detailed investigation of protein-protein interactions in the context of HIV-1 entry. Peptides are versatile tools to explore these interactions. For this purpose, peptides derived from a viral HPgV-1 protein and a combinatorial library approach should be used.

The first approach should be the investigation of a protein-derived peptide, called PEGI, and its structural prerequisites for its HIV-1 inhibitory activity as well as the breadth of HIV-1 inhibition. The PEGI peptide is derived from the E2 protein of the human Pegivirus 1 (HPgV-1). The HPgV-1 virus is known to interfere with the HIV-1 infection, but the details about the molecular mechanisms were still not completely understood. Using the PEGI lead peptide, a range of variants were planned to be synthesized to test their interference with HIV-1 infection. In addition, the mechanism of the HIV-1 inhibition should be determined on the protein level to characterize a potential binding site. The goal was to provide a molecular mechanism of the reported viral interference between HPgV-1 and HIV-1. Moreover, the relevance and mechanism of the HIV-1 inhibition by the protein-derived peptide (PEGI) should be studied with a HIV-1 strain panel and by cultivation of a PEGI-resistant HIV-1 isolate.

The second approach should be to screen broadly HIV-1 neutralizing antibody assays with synthetic combinatorial peptide libraries. Broadly HIV-1 neutralizing antibodies, targeting the envelope protein (ENV), are very potent HIV-1 inhibitors. Each antibody uses an individual mechanism for HIV-1 neutralization. The aim was to investigate if competitive assays, which are based on the binding of HIV-1 neutralizing antibodies to their target protein, can be used to identify new inhibitory peptides. The screening method should use a positional scanning synthetic combinatorial library (PS-SCL) based on a peptide scaffold, which is composed of positional scanning mixtures. The procedure should comprise these steps: scaffold ranking, positional scanning of the selected library and deconvolution and testing of the individual peptides.

3 Materials and Methods

3.1 Materials

3.1.1 Highly Purified water

The highly purified water (HPW) was provided by a Millipore water filtration system and was used for the preparation of all buffers and solutions if no other solvent was specified.

3.1.2 Chemicals

Employing the Fmoc-/tBu- (Fluorenyl-9-methyloxycarbonyl / *tert*-butyl) protecting group strategy the following building blocks were used for solid phase peptide synthesis (SPPS).

Table 2: Fmoc-building blocks.

Fmoc-building blocks	Distributor	Catalog Number
Fmoc-Ahx-OH	Iris Biotech	FAA1579.0100
Fmoc-D-Arg(Pbf)-OH	Senn Chemicals	02010
Fmoc-D-Asp(OtBu)-OH	Senn Chemicals	02199
Fmoc-D-Cys(Trt)-OH	MultiSynTech	1010
Fmoc-D-Glu(OtBu)-OH*H_2O	MultiSynTech	1051
Fmoc-D-Leu-OH	MultiSynTech	1025
Fmoc-D-Lys(Boc)-OH	Senn Chemicals	02022
Fmoc-D-Met-OH	MultiSynTech	1003
Fmoc-D-Phe-OH	MultiSynTech	1030
Fmoc-D-Thr(tBu)-OH	MultiSynTech	1053
Fmoc-D-Trp(tBu)-OH	Iris Biotech	FAA1339.0005
Fmoc-D-Tyr(tBu)-OH	MultiSynTech	1037
Fmoc-D-Val-OH	MultiSynTech	1039
Fmoc-L-Ala-OH*H_2O	Iris Biotech	FAA1005.0100
Fmoc-L-Arg(Pbf)-OH	Iris-Biotech	FAA1010.0100
Fmoc-L-Asn(Trt)-OH	Iris Biotech	FAA1015.0100
Fmoc-L-Asp(OtBu)-OH	Iris Biotech	FAA1020.0100
Fmoc-L-Cys(Trt)-OH	Iris Biotech	FAA1040.0100

Fmoc-L-Dap(Alloc)-OH	Neosystem	FA04014
Fmoc-L-Gln(Trt)-OH	Iris Biotech	FAA1043.0100
Fmoc-L-Glu(OAll)-OH	Neosystem	FA00601
Fmoc-L-Glu(OtBu)-OH*H_2O	Iris Biotech	FAA1045.0100
Fmoc-L-Gly-OH	Iris Biotech	FAA1050.0100
Fmoc-L-HCys(Trt)-OH	Iris Biotech	FAA1602 02
Fmoc-L-His(Trt)-OH	Iris Biotech	FAA1090.0100
Fmoc-L-Ile-OH	Iris Biotech	FAA1110.0100
Fmoc-L-Leu-OH	Iris Biotech	FAA1120.0100
Fmoc-L-Lys(Alloc)-OH	Neosystem	FA01203
Fmoc-L-Lys(Boc)-OH	Iris Biotech	FAA1125.0100
Fmoc-L-Met-OH	Iris Biotech	71989281.0100
Fmoc-L-Phe-OH	Iris Biotech	FAA1175.0250
Fmoc-L-Pro-OH	Iris Biotech	FAA1185.0100
Fmoc-L-Ser(tBu)-OH	Iris Biotech	FAA1190.0100
Fmoc-L-Thr(tBu)-OH	Iris Biotech	FAA1210.0100
Fmoc-L-Trp(tBu)-OH	Iris Biotech	FAA1225.0100
Fmoc-L-Tyr(tBu)-OH	Iris Biotech	FAA1230.0100
Fmoc-L-Val-OH	Iris Biotech	FAA1245.0100
Fmoc-O2Oc-OH (AOA)	Iris Biotech	FAA1435.0025

The following chemicals that were used for SPPS and immunoassays had a purity of at least 98%.

Table 3: Chemicals.

Chemicals	Distributor	Catalog Number
1,2-Ethandithiol	Sigma-Aldrich	
1,3-Dimethylbarbituric acid	Fluka	39565
2,2,2-Trifluoroethanol	Sigma-Aldrich	T6,300-2
Acetic acid	Sigma-Aldrich	1.00063.2500
Acetic anhydride	Roth	CP28.2
Acetonitrile (HPLC grade)	Fisher Chemical	
Ammonium acetate	Sigma-Aldrich	
Argon (g)	Linde	

Biotin	aMReSCO	0340-5G
Cyclohexane	Roth	6570.3
Dichlormethane (HPLC grade)	Fisher Chemical	D/1856/17
Dimethyl sulfoxide	Sigma-Aldrich	D8418
N,N-Diisopropylethylamine (DIPEA)	Sigma-Aldrich	496219-100mL
5,5'-Disulfanediylbis(2-nitrobenzoic acid) (DTNB)	Merck	1.03291.0005
G418 Geneticin (100 mg/Ml)	Invivogen	
Guanidinium hydrochloride (GuHCl)	Roth	6069.3
H-Asn-2-ClTrt resin (0.44 mmol/g)	Novabiochem	8.56140
H-Val-2-ClTrt resin (0.8 mmol/g)	Iris	RAA1301
Hydrochloric acid fuming 37 %	Roth	4625.2
Hydrogen peroxide solution 30 %	Sigma-Aldrich	H1009
Hygromycin B solution (50 mg/mL)	Sigma-Aldrich	
Methanol (HPLC grade)	Fisher Chemical	M4056/17
N,N-Diisopropylcarbodiimide	Iris-Biotech	RL-1015.0500
N,N-Dimethylformamide	Honeywell	
N,N-Dimethylformamide	Sigma-Aldrich	
N-Hydroxybenzotriazole	Sigma-Aldrich	
2,2-Dihydroxy-1H-indene-1,3(2H)-dione (Ninhydrin)	Pierce	21000
Nitrogen (g)	Linde	
o-Phenylenediamine dihydrochloride (OPD)	Sigma-Aldrich	P8287-100TAB
Ethyl cyano(hydroxyimino)acetate (Oxyma pure)	Iris Biotech	RL-1180.0500
Phenol	Roth	0040.1
Piperidine	Sigma-Aldrich	80640-1L
Potassium chloride	Fluka	60130
Potassium cyanide	Sigma-Aldrich	60187
Potassium dihydrogenphosphate	Roth	P749.3
Potassium hydrogenphosphate	Roth	P749.3
PS-2-CT-Gly resin (1.11 mmol/g)	Rapp Polymere	H751503313
Puromycin (5mg/ml)	SERVA Electrophoresis GmbH	

(Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyBOP)	Novabiochem	S5198009
Sodium acetate	Roth	6773.1
Sodium azide	Sigma-Aldrich	S2002
Sodium bicarbonate	Acros	424270010
Sodium carbonate	Roth	A135.2
Sodium chloride	Roth	3957.1
Sodium diethyldithiocarbamate trihydrate	Sigma-Aldrich	22,868-0
Sodium dihydrogenphosphate	Sigma-Aldrich	S0751
Sodium hydrogenphosphate	Sigma-Aldrich	S0876-1kg
Sulfuric acid 96 %.	Sigma-Aldrich	258105-1l
TentaGel S PHB-Gly resin (0.24 mmol/g)	Rapp Polymere	SA 1313
Tentagel S PHB-Trp resin (0.24mmol/g)	Rapp Polymere	SA1325
Tentagel S Ram resin (0.24mmol/g)	Rapp Polymere	S30 23
Tert-butyl-metyl ether	Merck	1.01843.2500
Tetrakis(triphenylphosphine)palladium(0) ($Pd(PPhe_3)_4$)	Sigma-Aldrich	216666-5G
Thioanisole	Fluka	101444882
4-(3,6-dimethyl-1,3-benzothiazol-3-ium-2-yl)-N,N-dimethylaniline chloride (Thioflavin T)	Sigma-Aldrich	T3516-5G
Trifluoroacetic acid, conc.	Roth	P088.3
Triisopropylsilane	Sigma-Aldrich	233781-50G
Tris(2-carboxyethyl)phosphine (TCEP)	Sigma-Aldrich	C4706-2g
2-Amino-2-(hydroxymethyl)-1,3-propanediol hydrochloride (Tris-HCl)	Roth	9090.3
2-Amino-2-(hydroxymethyl)-1,3-propanediol (Trizma Base)	Sigma-Aldrich	T6066-500G
Polyoxyethylene glycol sorbitan monolaurate (Tween 20)	Roth	9127.1

3.1.3 Buffers, Medium and Kits

Table 4: Buffers.

Name	Preparation
Assay coating buffer pH 9.6	4.0 mL 1 M Na_2CO_3 (aq) 8.0 mL 1 M $NaHCO_3$ (aq) diluted to 12 mL with HPW
10x Phosphate-buffered saline (PBS)	dissolved in 950 mL HPW: 14.2 g Na_2HPO_4 2.45 g KH_2PO_4 80.07 g NaCl 2.01 g KCl adjusted pH to 7.4 with HCl, dil. and filled up with HPW until 1000 mL
1x PBS	one to ten dilution of 10x PBS with HPW
Phosphate buffer pH 7.2 (PB)	71.7 mL 1 M KH_2PO_4 (aq) 28.3 mL 1 M K_2HPO_4 (aq) diluted to 1000 mL with HPW
Assay washing buffer	added 0.01 % (w/v) Tween 20 to phosphate buffer pH 7.2

Assay blocking buffer	added 1 % BSA (w/v) 0.01 % (w/v) Tween 20 to phosphate buffer pH 7.2
Assay main buffer	added 0.1 % BSA (w/v) 0.01 % (w/v) Tween 20 to phosphate buffer pH 7.2

Table 5: Medium.

Name	Additives	Distributor
RPMI 1640	10% heat inactivated fetal bovine serum (FBS), 0.3 mg/mL L-glutamine, 200 U/mL penicillin, 90 U/mL streptomycin	Pan Biotech GmbH

Table 6: Kits.

Name	Content	Distributor
Tropix Phospha-Light™ System	5x Dilution Buffer, Assay Buffer, CSPD®-Substrate, Reaction Buffer Diluent, Control Enzyme	Applied Biosystems™
ONE-Glo™ Luciferase Assay	ONE-Glo™ Luciferase Assay Buffer, ONE-Glo™ Luciferase Assay Substrate (lyophilized)	Promega Corporation

3.1.4 Proteins and antibodies

The recombinant proteins and antibodies were stored in aliquots in 1x PBS at -80 °C. Bovine serum albumin powder was stored at 4 °C. Streptavidin was stored in aliquots at 1 mg/mL in HPW at -20 °C.

Table 7: Proteins and antibodies.

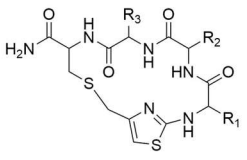
Proteins/ Antibodies	Distributor	Catalog number
447-52D	Polymun Scientific	AB014
Affinity purified anti-HIV-1-gp120 D7324 (Anti-gp120 D7324)	Aalto Bio Reagents	D7324
Anti-Fluorescein-HRP (Anti-Fluo-HRP)	Antibodies-online	ABIN237254
Anti-HIV-1 gp120 Monoclonal (2G12)	NIH AIDS Reagent Program	1476
Anti-HIV-1 gp120 Monoclonal (3BNC117)	NIH AIDS Reagent Program	12474
Anti-HIV-1 gp120 Monoclonal (VRC01)	NIH AIDS Reagent Program	12033
Anti-HIV-1 gp120 Monoclonal (VRC03)	NIH AIDS Reagent Program	12032
Anti-HIV-1 gp41 Monoclonal (F240)	NIH AIDS Reagent Program	7623
Anti-HIV-1 LAI gp120 Monoclonal (5F7)	NIH AIDS Reagent Program	2533
Anti-Human IgG (Fc specific) (Anti-Human)	Sigma-Aldrich	I-2136
Anti-Human IgG-Perox. (Anti-Human-HRP)	Sigma-Aldrich	A0170
Anti-mouse-HRP	Calbiochem	401207
b12	Polymun Scientific	AB011
Bovine serum albumin (BSA)	Sigma-Aldrich	A7906-50g
gp120(HXBc2) (Clade B)	Immune Tech	IT-001-0022p
gp41 (HXBc2)	Immune Tech	IT-001-005p

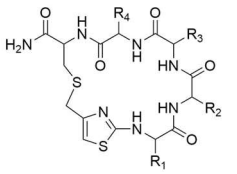
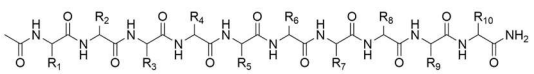
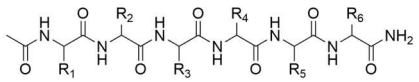
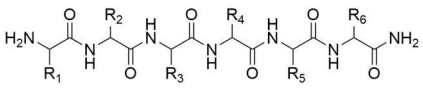
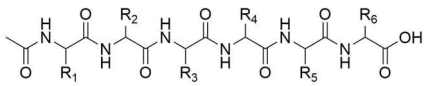
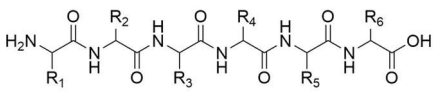
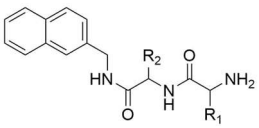
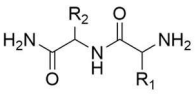
HIV-1 gp120 Monoclonal Antibody (F425 B4e8)	NIH AIDS Reagent Program	7626
HIV-1 gp120 Monoclonal Antibody (ID6)	NIH AIDS Reagent Program	2343
MAb X5	Dennis Burton Laboratory	
Monoclonal Anti-polyHistidine-Peroxidase antibody produced in mouse (Anti-His-HRP)	Sigma-Aldrich	A7058
Rabbit IgG anti-Sheep IgG (Fc)-HRP (Anti-sheep-HRP)	Dianova	313-035-046
Streptavidin	Thermo Fisher	T1015

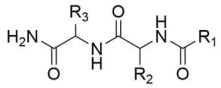
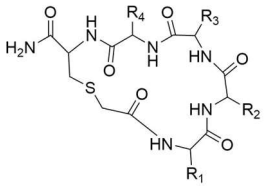
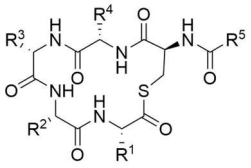
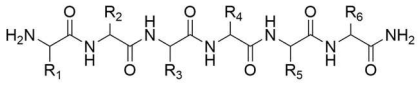
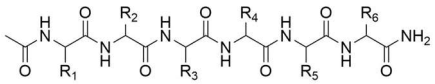
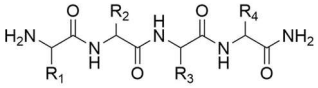
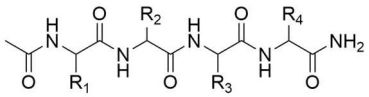
3.1.5 Combinatorial Libraries

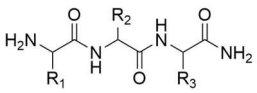
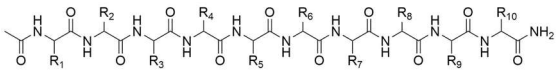
The following scaffold ranking samples (*all X* samples) were kindly provided by Torrey Pines Institute for Molecular Studies (TPIMS, Florida). The samples were dissolved in HPW with a final concentration of 2.5 mg/mL. The numbering is based on a four-digit code and the libraries are called TPI.

Table 8: Scaffold samples of the TPI libraries.

TPI	Number of Mixtures	Compounds per mixture	Total number of compounds	Structures
1988	188	2209	103823	Cyclic tripeptide thiazole 

1989	188	103823	4879681	Cyclic tetrapeptide thiazole 
2040	200	3.23 x 1011	4 x 1012	Decapeptide L-AA, acetlyated, amide 
2068	120	3200000	64000000	Hexapeptide (L-AA) acetylated, amide 
2069	120	3200000	64000000	Hexapeptide (L-AA) free amine, amide 
2085	120	3200000	64000000	Hexapeptide (L-AA) acetylated, carboxyl 
2119	120	3200000	64000000	Hexapeptide (L-AA) free amine, carboxyl 
2225	124	62	3844	N-naphthyl-dipeptide 
2226	124	62	3844	Dipeptide 

2228	228	3844-6448	399776	Acylated dipeptide 
2320	180	91125	4100625	Cyclic thio tetra peptide 
2321	86	68590-130321	1303210	Cyclic peptide 
2337	120	3200000	64000000	Hexapeptide (D-AA) free amine, amide 
2338	120	3200000	64000000	Hexapeptide (D-AA) acetylated, amide 
2339	120	8000	160000	Tetrapeptide (D-AA) free amine, amide 
2340	120	8000	160000	Tetrapeptide (D-AA) acetylated, amide 

2500	Tripeptide (L-, D-AA), free amine, amide 
2512 210	Decapeptide w/citrulline L-AA, acetylated, amide 

3.1.6 Consumables

Table 9: Consumables.

Consumable	Distributor
2 mL reaction tubes	Sarstedt
5 mL tubes	Nerbe
15 mL Falcon tubes	Sarstedt/ Corning
50 mL Falcon tubes	Sarstedt/ Corning
Immunolon 2HB Plates 96 well	Thermo Scientific
Microplate, 96well, PS, F-Bottom, Black, non-binding	Greiner-Bio One
Round bottom 96 well plate, natural PP, non-sterile	Fisher Scientific
5 mL Syringe	B. Braun
UV-VIS cuvettes	Brand
Scintillation vials	Wheaton
Nunc™ F96 Microwell™ plates	Thermo Scientific
Cell culture plates 24wells and 96 wells	Sarstedt

Cell culture flasks	Sarstedt
Serological pipettes	Nerbe
Pipette tips with filter, Biosphere	Sarstedt
0.5 mL reaction tubes	Sarstedt
1.5 mL reaction tubes	Sarstedt
Pipette tips (1-10 mL)	Brand
Pipette tips (0.5-10 µL, 20-200 µL, 100-1000 µL)	Eppendorf
0.22 µm syringe filters	Roth

3.1.7 Machines

Table 10: Machines.

Machine	Manufacturer
Liquid handler:	
Precision	BioTek
Plate washers:	
ElkX405 (automatic)	BioTek
CappWash (manual)	CAPP
LC-MS:	
Agilent 1100	HP (Hewlett Packard)
(Degasser, Binary Pump, Column Oven, DAD detector)	
Autosampler Series 2000	Perkin Elmer
Column: Kinetex C-18 (2.6 µM, 100 Å; 50 mm x 2.1 mm)	Phenomenex
Mass spectrometer: API 2000, ESI source	AppliedBiosystem
Plate reader:	
Infinite Reader F200	Tecan
Spectra Max	Molecular Devices
Victor³ V 1420	Perkin Elmer

NanoPhotometer	Implen
Freeze dryer: ALPHA 1-4:	 Christ
pH-Meter: Seven Easy	 Mettler Toledo
Preparative HPLCs: Pumps: L-6200	
	Merck Hitachi,
	Merck Hitachi
UV-VIS detectors:	
L-4250, Variable Wavelength	Knauer
Pharmacia LKB GE UV-1	Pharmacia
Waters 486	Waters
Columns:	
Kinetex C18 (5 µM, 100 Å, 100 x 21.2 mm)	Phenomenex
Reprosil C18 (5 µM, 100 Å, 250 x 25 mm)	Dr. Maisch
Fraction collectors:	
SuperFrac	GE Healthcare
Synthesis robots:	
Syro 1	MultiSyntech
Respep	Intavis
Incubators:	
Innova 42	New Brunswick
HERACELL 150i CO2	Thermo Fisher
HERAEUS 6000 Inkubator Typ BB 6220 CU	Heraus
Freezer (-80°C): Innova ULT	 New Brunswick

Shakers:	
Titramx 101	Heidolph
VTX-3000L	LMS Harmony
PS M3D	Grant Instruments
Centrifuges:	
Galaxy Mini Star	VWR
Galaxy 7D	VWR
Megafuge 1.OR	Hereaus
Sepatec	Hereaus
Rotina 380 R	Hettich
5810 R	Eppendorf
Water filtration System:	
Milipore Synergy 185 (filter: SymPak2)	Merck
Balances:	
CPA225D	Sartorius
BP210S	Sartorius
Thermomixer Comfort	Eppendorf

3.2 Peptide Synthesis

All peptides presented in this work were synthesized *via* a solid phase peptide synthesis (SPPS) approach in a 25 µmol batch. The synthesis advanced from the C-Terminus to the N-Terminus of the peptide sequence. The synthesis was based on the Fmoc/tbu protecting group strategy [4].

It was conducted in 2.0 mL syringes equipped with a frit in the following called reactor. To start every synthesis batch 100 mg resin per peptide to be synthesized was weighed into a reactor and swollen in DMF for at least 20 min. Depending on the C-Terminus of the peptide sequence, different resins were used (Table 11).

Table 11: Synthesized peptides sorted by resin.

Rink amide resin	C-terminal amino acid preloaded PHB or Wang resin
PEGI-D2, D3, D6, D7, D31, D32, D34-D36, D49-68	PEGI, PEGI-Biotin
Disulfide $Loop_{HXBc2}$-ox	PEGI-D4, D5, D8-D30, D33, D37-D48, D51-D60
VRC01-crude-1-24, VRC01-ReH2-17, VRC01-ReH11-Bio, VRC01-ReH12-Bio, VRC01-ReH13-Bio	
V3-$Loop_{HXBc2}$	

3.2.1 Automated SPPS

The key principle of SPPS is the coupling of amino acid building blocks in a repetitive cycle of thefollowing steps: N-terminal Fmoc-deprotection, coupling and capping. The synthesis was mostly achieved with automated pipetting robots using a standardized pipetting procedure. Each Fmoc-amino acid was coupled using a repetitive cycle-based synthesis. The cycle varied for the coupling of D- and L- amino acids (in regard of the equivalents, coupling time and replicates used). Fmoc-O2Oc-OH was coupled with the D-amino acid and Fmoc-Ahx-OH with the L-amino acid protocol. The details of the synthesis cycle are shown in Table 12. Due to sequence dependent difficult couplings, solubility issue or high-cost Fmoc-building blocks were excluded from the automated synthesis and coupled by manual protocols (cf. 3.2.2.1).

Table 12: Automated synthesis cycle.

Step	Procedure	Time [min]	Volume [µL]	Reagents	Replicates
1	1. Fmoc-deprotection	5	500	20 % (v/v) Piperidine in DMF	1
	2. Fmoc-deprotection	15			1
	Washing	1	800	DMF	5
2	L-AA Coupling	60	350	0.33 M Fmoc-AA	2
			150	20 % DIC (v/v) in DMF	
	D-AA Coupling	120	210	0.34 M Fmoc AA and 0.5 M Oxyma Pure in DMF	1
			200	0.5 M Oxyma DMF in DMF	
			90	20 % DIC (v/v) in DMF	
			2	DMF	
	Washing	1	800	DMF	5
3	Capping	20	500	33.3 % (v/v) Pyridine, 16.7 % (v/v) Acetanhydride in DMF	1
	Washing	1	800	DMF	5

3.2.2 Non-automated SPPS Procedures

3.2.2.1 Manual Couplings

The manual coupling followed the same steps (Fmoc-deprotection, Coupling, Capping) as the automated synthesis. Building blocks with orthogonal protecting groups, D-AAs and sequence dependent difficult couplings were coupled manually for the successful synthesis of PEGI

peptides. Biotin was always coupled manually using the same protocol due to its solubility. The resin in the reactor was swollen at least 20 min with DMF before any reaction was performed.

For Fmoc-deprotection the resin was incubated with piperidine in DMF (500 µL 20 % (v/v)) on the shaker for 30 min. This was followed by washing five times with DMF (800 µL). Manual couplings were incubated overnight on the shaker. The coupling reagent was DIC (1.5 eq. per eq. building block) in Oxyma Pure / DMF (500 µL 0.5 M). The specific conditions for each building block are listed in Table 13.

Table 13: Conditions for the manual coupling.

Building block	eq.	Coupling time	eq. DIC	4M GuHCl	
Biotin	2	overnight	3	x	in 500 µL 0.5 M Oxyma Pure/DMF
Fmoc-D-Cys(Trt)-OH	3		4.5	x	
Fmoc-D-Asp(OtBu)-OH	3		4.5	x	
Fmoc-D-Trp(tBu)-OH	3		4.5	x	
Fmoc-L-HCys(Trt)-OH	5		7.5	x	
Fmoc-L-Val-OH	7		10.5	x	
Fmoc-L-Asn(Trt)-OH	7		10.5	x	
Fmoc-L-Lys(Alloc)-OH	5		7.5		
Fmoc-L-Glu(OAll)-OH	5		7.5		
Fmoc-L-Dap(Alloc)-OH	5		7.5		

The chaotropic salt GuHCl was added to the reaction to prevent incomplete coupling due to aggregation on the resin [96]. This was needed for the synthesis for all PEGI derivates at position 50 and 56 (Figure 10). After the overnight incubations the reactors were washed five times with 800 µL DMF.

$_{45}$LCDCPNGPWVWVPAFCQAVG$_{64}$

Figure 10: Lead structure of PEGI peptides, Positions coupled with addition of GuHCl are highlighted in blue.

The manual coupling was followed by the capping step. The reactor with the resin was incubated with pyridine (33.3 % (v/v)), acetanhydride (16.7 % (v/v)) in DMF (600 µL) on the shaker for 20 min. This was followed by washing five times with DMF (800 µL).

3.2.2.2 N-Terminus Acetylation

The N-terminal amine of the synthesized peptide chain can be modified. In this work it was acetylated manually by the following protocol. The resin in the reactor was swollen at least 20 min with DMF. The first step was Fmoc-deprotection. The resin was incubated with piperidine (20 % (v/v)) in DMF (500 µL) on the shaker for 30 min. It was followed by washing five times with 800 µL DMF. The second step was the incubation with pyridine (33.3 % (v/v)), acetanhydride (16.7 % (v/v)) in DMF (600 µL) on the shaker for 20 min. Finally, the reactor with the resin was washed five times with DMF (800 µL).

3.2.2.3 On Resin Lactam Formation

For the lactam bridge formation of PEGI-D33, -D47 and -D48, a resin based approach was used. The peptides were synthesized with orthogonal protected Fmoc-building blocks. Before the lactam bridge formation, the side chain protecting groups were cleaved selectively by the following protocol: 1,3-Dimethylbarbituric acid (19.3 mg) and $Pd(PPhe_3)_4$ (12.9 mg) were dissolved in argon-sparged DMF (1.5 mL). The mixture was drawn up in the reactor and agitated constantly in the dark for 3 h. After incubation the resin in the reactor was washed with DCM (two replicates), DMF (two replicates), 0.5 % (v/v) Sodium diethyldithiocarbamate trihydrate in DMF (eight replicates) DMF (three replicates) and DCM (two replicates) (800 µL each). For the lactam bridge formation, the resin was incubated with PyBOP (3 eq.) and *N,N*-Diisoproplyethylamine (DIPEA) (6 eq.) in DMF (400 µL) on the shaker overnight. After the incubation the resin was washed five times with DMF (800 µL).

3.2.2.4 *Kaiser* Test

Coupling reactions during SPPS can be monitored with a color reaction called *Kaiser* Test. The test is a sensitive method to detect free amines *via* the ninhydrin reaction [97]. The resin was washed with 800 µL DMF (3 replicates), DCM (3 replicates) and ethanol (1 replicate). A small

amount of resin was taken into a test tube. The following solutions were added to the test tube: KCN in pyridine (2-3 drops 0.07 mg/mL), ninhydrin in ethanol (2-3 drops 0.05 mg/mL) and phenol in ethanol (2-3 drops 2 mg/mL). The mixture was heated up to 95 °C under constant shaking for 5 minutes. If the test was positive a blue color was observed. Consequently, the preceding incomplete coupling reaction was repeated after a positive *Kaiser* test.

3.2.2.5 Disulfide Bridge Formation

Free cysteines can be oxidized by intramolecular disulfide bridge formation. During this work peptides PEGI-D50, V3-$Loop_{HXBc2}$ and Disulfide $Loop_{HXBc2}$-ox were oxidized by atmospheric oxygen in ammonium acetate buffer. Peptides were dissolved at 0.3 mg/mL in oxidation buffer a mixture of acetonitrile and 0.1 M ammonium acetate (aq, pH 8) in a ratio of 50:50. This procedure was performed subsequently after total cleavage and first purification by preparative HPLC. The dissolved peptides were shaken constantly at room temperature for 3 days. The reaction was monitored with a method called *Ellmans* test. It is a color reaction based method to determine free cysteine [98, 99]. The test was performed by a standard protocol. Briefly, DTNB in 50 mM sodium acetate (50 µL 2 mM)), Tris (100 µL 50 mM), HPW (80 µL) and a sample of peptide solution at 0.3 mg/mL in oxidation buffer (10 µL) were mixed. After 5 min the absorption was measured in a cuvette at 412 nm. After the endpoint of the disulfide bridge formation was verified with a negative *Ellmans* test the peptide solution was frozen, lyophilized and purified *via* HPLC for a second time.

3.2.2.6 Total Cleavage

The cleavage of side chain protection and C-terminal resin linkage to liberate the peptide after synthesis was achieved in a simultaneous approach called total cleavage. At first the resin inside reactors was washed with DMF (5 replicates) and DCM (2 replicates) (800 µL each). The washed resin was incubated with cleavage mixture (1.5 mL) under constant shaking for 180 min and a second time with cleavage mixture (1.0 mL) for 60 min. The cleavage mixture contained TFA (82.5 % (v/v)), phenol (5 % (w/v)), HPW (5 % (v/v)), thioanisole (5 % (v/v)) and 1,2-ethandithiol (2.5 % (v/v)). The incubation was finished by precipitating cleaved peptide containing cleavage mixture into a scintillation vial filled with 10 mL cold cyclohexane/ *tert*-butyl-methyl ether (50/50). 10 mL HPW was added to the scintillation vial. The vial was vortexed for 30 s and

centrifugated at 700 g for 10 min. This was followed by freezing, removal of the cyclohexane/ *tert*-Butyl-metyl ether (50/50) phase and lyophilization of aqueous layer containing the crude peptide. The resulting powder was dissolved (10 mL ACN/HPW (50:50) + 0.1 % TFA) and lyophilized twice. It was analyzed by LC-MS method synthesis control (cf. 3.2.4).

3.2.2.7 Test Cleavage

For monitoring SPPS at critical coupling steps and before the total cleavage, a test cleavage was performed. Initially, the resin was washed with 800 µL DMF (3 replicates) and DCM (2 replicates). A small amount of resin sample was transferred to a test tube and 100 µL test cleavage mixture was added. It was incubated for 90 min. The test cleavage mixture contained 70 % (v/v) TFA, 20 % (v/v) DCM, 5 % (v/v) HPW and 5 % (v/v) TIPS. After the incubation the resin sample was dried under nitrogen flow. The dried resin sample was suspended in 100 µL ACN/HPW (50/50) + 0.1 % TFA, vortexed for 30 s and centrifugated at 2600 g for 30 seconds. 40 µl of the supernatant were used for the analysis with LC-MS. LC-MS method synthesis control was applied (cf. 3.2.4).

3.2.3 Preparative HPLC

After total cleavage the peptides were purified via RP-HPLC. The purifications were performed on C18-columns with a flow rate of 3 or 9 mL/min and a fraction for each minute of the run was collected separately. For PEGI peptides a mixture of TFE/AcOH/HPW was used as injection solution. All other peptides were dissolved in TFA prior to injection. Oxidized peptides were dissolved in acetonitrile/HPW mixture corresponding to the starting conditions of the chosen gradient for purification (cf. 3.2.2.5). The gradients for each peptide purification are listed in Table 14.

Table 14: Preparative HPLC methods used for peptide purification.

	Peptides	Gradient percentage ACN	Duration [min]	Flowrate [mL/min]
1. Purification	PEGI. PEGI-Biotin, PEGI-D2-D68, Disulfide $Loop_{HXBc2}$-ox	20 % -60 %	100	9
	F240-ReH1-ReH9, F240-ReH1-Bio, F240-ReH2-Bio VRC01-ReH2-ReH17, VRC01-ReH11 Bio, VRC01-ReH12-Bio, VRC01-ReH13-Bio	10 % - 50 %	80	
2.Purification after oxidation	Disulfide $Loop_{HXBc2}$-ox	20 % - 55 %	80	3
	V3-$Loop_{HXBc2}$	10 % - 50 %	80	
	PEGI-D50	20 % - 60 %	60	

All purifications were monitored at wavelengths of 220 nm and 280 nm. Based on the obtained UV-chromatograms 40 µl samples of the fraction of interest were taken and used for LC-MS measurement (method purification) (cf. 3.2.4). Fractions containing peptides with a certain level of purity were pooled together, frozen, lyophilized and final LC-MS analysis was performed (method final product). The analyzed final product data for the synthesized peptide can be found in section 7.1. The level of purity of the peptides was determined based on the UV chromatogram. Purity was determined by UV peak purity and only fractions with main peak purity of 90-95 % were accepted. The standard procedure was to allocate mass charge ratio (m/z) signals of the ESI- spectrum (Q1) over noise level to molecular mass of the analyzed peptide. The ESI spectrum time always correlated with the UV main peak retention time.

3.2.4 LC-MS Analysis

The peptides were analyzed *via* LC-MS during the synthesis, as well as, after cleavage and purification. It was also used for the final quality control of purified peptide (c.f. 3.2.3). Following parameters were constant for all measurement methods: mobile phase components (ACN/HPW + 0.1 % TFA), solvent of the sample (ACN/HPW + 0.1 % TFA), HPLC column, column temperature, ESI source and UV detection at 220 nm. Important criteria for the analysis were the purity of the UV chromatogram and allocation of mass to charge ratio (m/z) of ESI mass spectrum with the molecular mass of peptide sequence. Routinely the retention time of ESI mass spectrum matching UV main peak was analyzed for m/z coinciding with the molecular weight of the peptide of interest. It should be noted that more than one single mass signal corresponding to the UV main peak was detected for the VRC01-ReH peptides. The used LC-MS methods are shown in Table 15. The analyzed final product data for the synthesized peptide can be found in in section 7.1.

Table 15: LC-MS methods for peptide analysis

LC-MS method	Injection volume	Gradient
purification	20 µL	5 % - 95 % ACN in 5 min
synthesis control		5 % - 95 % ACN in 10 min
final product		5 % - 80 % ACN in 15 min

3.2.5 Storage and Stock Solutions

Lyophilized peptides were stored at -20 °C for long term storage. Peptide stock solutions were stored at -4 °C. DMSO and ACN/HPW (50:50) were used as stock solution solvents. Stock solutions were prepared depending on the purpose at four different concentrations (2.5/10/12/13 mM).

3.3 Enzyme linked Immunosorbent Assays (ELISAs)

Enzyme linked immunosorbent assays (ELISAs) are a robust method to investigate protein-peptide interaction in binding and competitive assay set ups. If not specifically mentioned, the following materials, reagents and conditions were used as standard protocol. ELISAs were performed on Immunolon 2HB 96 well plates. The following steps were performed for standard incubation cycles on 96 well plates: Coating of protein or compound of interest in Assay coating buffer pH 9.6, blocking with assay blocking buffer, assay format depending number of incubation steps of protein or peptide of interest in assay main buffer, assay format depending number of incubation steps of detection antibody and a final detection with a color reaction. Each incubation step was followed by washing four times with assay washing buffer (300 µL/well). Detection antibodies were labeled with horse-radish peroxidase (HRP). The enzymatic color reaction was conducted with o-Phenylenediamine dihydrochloride (OPD) (1mg/mL) + 0.03 % (v/v) H_2O_2 (100 µL/well). After 20 min the color reaction was stopped with H_2SO_4 (50 µL/well, 2 M) and the plate was measured at a wavelength of 492 nm. For each experiment the target interaction was done in duplicates in addition to the corresponding blanks. Every experiment was subsequently repeated at least once (n_{min}= 4). Blanks were used as a quality control for unspecific signal effects and subtracted from the sample signal. Data is always presented as AVG ± SD.

3.3.1 Binding Assays

3.3.1.1 Binding Assay Format A

To investigate PEGI-and its derivates binding to $gp120_{HXBc2}$ and $gp41_{HXBc2}$ a streptavidin-based ELISA with a 3 h ligand incubation step was used. For the coating step streptavidin (100 µL/well,4 µg/mL) in assay coating buffer was incubated at 4 °C overnight. The next day the plate was blocked with assay blocking buffer (200 µL/well) at room temperature under constant shaking for 1 h. In the next step biotinylated peptides (100 µL/well, 5 µM) were incubated at room temperature under constant shaking for 2 h. The peptides were diluted from 2.5 or 10 mM stock solution with assay main puffer yielding a concentration of 5 µM. As ligand proteins $gp120_{HXBc2}$ or $gp41_{HXBc2}$ (100 µL/well) were incubated at three different doses

($gp120_{HXBc2}$: 1.5/0.75/0.38 µg/mL and $gp41_{HXBc2}$: 1.0/0.5/0.25 µg/mL) at room temperature under constant shaking for 3 h. For $gp41_{HXBc2}$ assays anti-His-HRP (100 µL/well, 1/5000 dilution) were incubated at room temperature under constant shaking for 1 h. For $gp120_{HXBc2}$ assays anti-gp120 D7324 (100 µL/well 1/5000 dilution) were incubated at room temperature under constant shaking for 1 h. Followed by the incubation of anti-sheep-HRP (100 µL/well, 1/5000 dilution) at room temperature under constant shaking for 1 h. The plates were washed and detection performed by the standard protocol. Blanks lacking the protein or the peptide of interest were included on each plate. Blanks lacking the protein were then subtracted from the sample signal for the data analysis. PEGI-Biotin was used as an internal standard for comparability of each plate and used for OD_{rel} calculation.

3.3.1.2 Binding Assay Format B

To investigate molecular binding of peptides and protein/antibodies of interest a streptavidin-based ELISA protocol with a ligand incubation step over night was used. For the coating step streptavidin (100 µL/well, 4 µg/mL) in assay coating buffer was incubated at 4 °C overnight. On the next day the plate was blocked with assay blocking buffer (200 µL/well) at 20 °C under constant shaking for 1 h. In the next step biotinylated peptides (100 µL/well, 5 µM) were incubated at 20 °C under constant shaking for 2 h. The peptides were diluted from 2.5 or 10 mM stock solution with assay main puffer yielding a concentration of 5 µM. Protein or antibody (100 µL/well) was incubated as the ligand in a two-fold serial dilution at 4 °C for 18 h. Concentrations that were used for each protein/antibody and the corresponding detection antibody or antibodies are listed in Table 16. For proteins with a single detection antibody the corresponding antibody-HRP (100 µL/well,1/5000 dilution) was incubated at 20 °C under constant shaking for 1 h. For proteins with two detection antibodies the first corresponding antibody (100 µL/well, 1/5000 dilution) was incubated at 20°C under constant shaking for 1 h and it was followed by the incubation of corresponding antibody-HRP (100 µL/well, 1/5000 dilution) at 20 °C under constant shaking for 1 h. The plates were washed and detection performed by the standard protocol. The incubation steps at 4 °C were conducted in a moisturized covered plastic box. Blanks lacking the protein or the peptide of interest were included on each plate. Blanks lacking the protein were then subtracted from the sample signal for the data analysis.

Table 16: Protein and antibody concentrations for Binding Assay Format B.

Protein/ antibody	Tested concentration [µg/mL or dilution ratio]	Detection antibody	Dilution ratio detection antibody
$gp120_{HXBc2}$	1.50	anti-gp120 D7324+ anti-sheep-HRP	1/5000
	0.75		
	0.38		
$gp41_{HXBc2}$	1.00	anti-His-HRP	
	0.50		
	0.25		
2G12	0.050	anti-human-HRP	
	0.025		
	0.013		
	0.006		
F425 Be8	1.000		
	0.500		
	0.250		
	0.125		
3BNC117	0.040		
	0.020		
	0.010		
	0.005		
VRC01	0.100		
	0.050		
	0.025		
	0.013		
447-52D	0.010		
	0.005		
	0.003		
	0.001		
F240	0.020		
	0.010		
	0.005		
	0.003		
b12	0.100		
	0.050		
	0.025		
	0.013		
X5	2.000		
	1.000		
	0.500		
	0.250		
5F7	1/200	anti-mouse-HRP	
	1/400		
	1/800		
	1/1600		

3.3.1.3 Binding Assay Format C

To investigate peptide-peptide interaction a streptavidin-based ELISA with an overnight incubation step was used. For the coating step streptavidin (100 µL/well 4 µg/mL) in assay coating buffer was incubated at 4 °C overnight. The next day the plate was blocked with assay blocking buffer (200 µL/well) at 20 °C under constant shaking for 1 h. In the next step biotinylated peptides (100 µL/well 5 µM) were incubated at room temperature under constant shaking for 2 h. The peptides were diluted from 2.5 or 10 mM stock solution with assay main puffer to a final concentration of 5 µM. As ligand 100 µL/well fluorescein tagged peptide was incubated in a two-fold serial dilution (starting 25 nm) at 4 °C for 18 h. This was followed by the incubation of 100 µL/well anti-Fluo-HRP (1/10000 dilution) at 20 °C under constant shaking for 1 h. The plates were washed and detection performed by the standard protocol. The incubation steps at 4 °C were conducted in a moisturized covered plastic box. A concentration dependent blank missing the biotinylated peptide was included and subtracted from the corresponding concentration of the sample.

3.3.2 Competitive Assay

3.3.2.1 Competitive Assay Format A

A competitive ELISA based on the F240-gp41$_{HXBc2}$-interaction was designed for the screening of combinatorial library samples. For the coating step anti-human antibody (100 µL/well, 1/2500 dilution) was incubated at 4 °C overnight. The plate was blocked with 200 µL/well assay blocking buffer at room temperature under constant shaking the following day for 1 h. The next step was the incubation of F240 antibody (100 µL/well, 0.016 µg/mL) at room temperature under constant shaking for 2.5 h. The combinatorial library samples and gp41$_{HXBc2}$ were pipetted successively to the wells. The library samples were prepared on 96 wells polypropylene plates (PP plates). The first transfer to the PP plates was pipetted manually. The second dilution step and pipetting library sample (50 µL/well) to the final assay plate was conducted with a liquid handler. The tested concentrations of the library samples are listed in Table 17. After the library sample was pipetted, gp41$_{HXBc2}$ (50 µL/well, 0.2 µg/mL) was manually added. The final volume per well was 100 µL. The plate was incubated at 4 °C for 18 h. On the next day

anti-His-HRP (100 µL/ well, 1/5000 dilution) was incubated at room temperature under constant shaking for 1 h. The detection was performed by the standard protocol. For the washing steps a plate washer was used with the same volumes and replicates as the standard protocol.

Table 17: Library sample concentrations for Competitive Assay Format A.

Type of library sample	Name	Conc used [µg/mL]	Final conc [µg/mL]	Solvent
Scaffold ranking	library set	4000	2000	HPW
		2500	1250	
		1250	625	
Positional scanning	TPI 2338	2000	1000	

The incubation steps at 4 °C were conducted in a moisturized covered plastic box. For 100 % binding controls the library sample was replaced with the sample solvent. As inhibition controls three doses (2.5/0.16/0.05 nM) Disulfide $Loop_{HXBc2}$-ox were included on every plate. Blanks lacking $gp41_{HXBc2}$ or F240 of interest were included additionally and the average blank subtracted from all signals.

3.3.2.2 Competitive Assay Format B

A competitive ELISA protocol based on antibody-$gp120_{HXBc2}$-interaction was developed for the screening of combinatorial library samples. Three $gp120_{HXBc2}$ targeting antibodies were tested: b12, VRC01 and VRC03. The first step was the coating of $gp120_{HXBc2}$ (100 µL/well 0.5 µg/mL) at room temperature under constant shaking for 2 h. This was followed by blocking with 200 µL/well assay blocking buffer at room temperature under constant shaking for 1 h. After the blocking step library samples and antibodies were pipetted to the plate. The library samples were prepared on 96 wells PP plates. The first liquid transfer to the PP plates was pipetted manually. The second dilution step and pipetting library sample (50 µL/well) to the final assay

plate were conducted with a liquid handler. The library samples and concentrations tested with each individual antibody are listed in Table 18. After the library sample (50 µL/well) was pipetted, b12 (50 µL/well 0.2 µg/mL), VRC01 (50 µL/well 0.2 µg/mL) or VRC03 (50 µL/well 5.0 µg/mL) was added manually. The final volume per well was 100 µL. The plate was incubated at 4 °C for 18 h. The next day anti-human-HRP (100 µL/well 1/5000 dilution) was incubated at room temperature under constant shaking for 1 h. The detection was performed by the standard protocol. For the washing steps a plate washer was used with the same volumes and replicates as the standard protocol. The incubation steps at 4 °C were conducted in a moisturized covered plastic box. For 100 % binding controls the peptide sample was replaced with assay main buffer. As inhibition control $gp120_{HXBc2}$ (2.5 µg/mL) was added instead of a library or peptide sample on every plate. Blanks lacking $gp120_{HXBc2}$ or the used antibody were included on all plates and the average blank subtracted from all signals.

Table 18: Library sample concentrations for Competitive Assay Format B.

Antibody	Type of library sample	Name	c used [µg/mL]	Final c [µg/mL]	Solvent
b12	Scaffold Ranking	library set	4000	2000	HPW
			2500	1250	
			1250	625	
	Positional Scanning	TPI 2338	625	313	
			313	157	
VRC03	Scaffold Ranking	library set	4000	2000	
			2500	1250	
			1250	625	
VRC01	Scaffold Ranking	library set	4000	2000	
			2500	1250	
			1250	625	
	Positional Scanning	TPI 2338	1250	625	
			625	313	

3.3.2.3 Competitive Assay Format C

The competitive ELISA based on the F240-$gp41_{HXBc2}$-interation was optimized for the testing of crude and purified peptides of interest. For the coating anti-human antibody (100 µL/well, 1/2500 dilution) was incubated at 4 °C overnight. The next day the plate was blocked with assay blocking buffer (200 µL/well) at 20 °C under constant shaking for 1 h. The following step was the incubation of F240 antibody (100 µL/well, 0.016 µg/mL) at 20 °C under constant shaking

for 2.5 h. This was followed by the successive pipetting of peptides and $gp41_{HXBc2}$ to the plate. Peptides were diluted with HPW or assay main buffer from 12 mM stock solution in DMSO in cluster tubes. Crude peptides were tested at two different concentrations and purified peptides in a serial four-fold dilution. 50 µL/well diluted peptide samples were transferred to each assay plate. After this $gp41_{HXBc2}$ (50 µL, 0.2 µg/mL) was added to each well. The final volume per well was 100 µL. The plate was incubated at 4 °C for 18 h. On the next day anti-HIS-HRP (100 µL/ well, 1/5000 dilution) was incubated at 20 °C under constant shaking for 1 h. The plates were washed and detection performed by the standard protocol. The incubation steps at 4 °C were conducted in a moisturized covered plastic box. For 100% binding controls the peptide sample was replaced with assay main buffer. As inhibition controls three concentrations (2.5/0.16/0.05 nM) Disulfide $Loop_{HXBc2}$-ox were included on every plate. Blanks lacking $gp41_{HXBc2}$ or F240 were included additionally and the average blank subtracted from all signals.

3.3.2.4 Competitive Assay Format D

The competitive ELISA protocol based on antibody-$gp120_{HXBc2}$-interaction was optimized for the testing of crude and purified peptides of interest with a panel of $gp120_{HXBc2}$ targeting antibodies. The first step was the coating of $gp120_{HXBc2}$ (100 µL/well, 0.5 µg/mL) at 20 °C under constant shaking for 2 h. This was followed by blocking the wells with assay blocking buffer (200 µL/well) at 20 °C under constant shaking for 1 h. The peptides to be tested were diluted with assay main buffer from stock solution in DMSO (10/12/13 mM) in cluster tubes. Crude peptides were tested at two concentrations with VRC01 and at one concentration with b12. Purified peptides were tested in a serial four-fold dilution. 50 µL/well of the diluted peptide were transferred to the plate and 50 µL/well of the antibody of interest was added. The final volume per well was 100 µL. The antibody panel and used concentrations are listed in Table 19. The plate was incubated at 4 °C for 18 h. On the next day anti-human-HRP (100 µL/well, 1/5000 dilution) was incubated at 20 °C under constant shaking for 1 h. The plates were washed and detection performed by the standard protocol. The incubation steps at 4 °C were conducted in a moisturized covered plastic box. For 100 % binding controls the peptide sample was replaced with assay main buffer. As inhibition control $gp120_{HXBc2}$ (2.5 µg/mL) was added

instead of a peptide sample on every plate. Blanks lacking $gp120_{HXBc2}$ or the used antibody were included on all plates and average blanks subtracted from all signals.

Table 19: Antibody panel concentrations for Competitive Assay Fomat D.

Antibody	Used concentration [µg/mL]	Final concentration per well [µg/mL]
2G12	0.1	0.05
F425 B4e8	3	1.5
3BNC117	0.04	0.02
VRC01	0.2	0.1
447-52D	0.02	0.01
5F7	1/100 dilution	1/200 dilution
ID6	1/500 dilution	1/1000 dilution
b12	0.2	0.1
X5	4	2

3.3.2.5 Competitive Assay Format E

A competitive streptavidin ELISA protocol based on F425 B4e8–V3-$Loop_{HXBc2}$ interaction was established. The first step was coating of streptavidin (100 µL/well, 4 µg/mL) at 4 °C overnight. The next day the plate was blocked with assay blocking buffer (200 µL/well) at 20 °C under constant shaking for 1 h. After this V3-$Loop_{HXBc2}$ (100 µL/well, 5 µM) was incubated at 20 °C under constant shaking for 2 h. The peptides were tested in a serial four-fold dilution and diluted from 10 mM stock solution with assay main buffer. 50 µL/well of the diluted peptide were transferred to the plate and 50 µL/well of F425 B4e8 (9.3 nM) were added. The final volume per well was 100 µL. The plate was incubated at 4 °C for 18 h. On the next day 100 µL/ well anti-human-HRP (1/5000 dilution) was incubated at 20 °C under constant shaking for 1 h. The plates were washed and detected with the standard protocol. For the incubation step at 4 °C the plates were placed in a plastic box covered with a lid. The box was moisturized with wet paper towels. For 100 % binding controls the peptide was replaced with Assay main buffer. As inhibition control 25 µM V3-$Loop_{HXBc2}$ was added instead of the tested peptide sample on every plate. Blanks lacking $gp120_{HXBc2}$ or the used antibody were included on all plates and average blank subtracted from all signals.

3.4 Thioflavin Assay

The Thioflavin T Assay is a fluorescence measurement method to determine peptide binding to the dye Thioflavin T [100-102]. The assay was conducted in black non-binding 96 well plates. Sterile filtered (0.22 µm filter) phosphate buffer pH 7.2 (PB) was used for this assay. Thioflavin stock solution in PB (0.8 mg/mL) was kept at 4 °C in the dark for long term storage. Thioflavin stock solution was diluted 1 to 50 with PB yielding Thioflavin working solution. Firstly, Thioflavin working solution (148 µL/well) was pipetted. 2.5 mM peptide stock solution (1.5 µL/well) was added and subsequently mixed five times manually. The peptide stock solutions were prepared with DMSO. For this assay higher concentrated stock solution were diluted to 2.5 mM with DMSO. The plate was incubated at room temperature under constant shaking for 1 h. After the incubation the fluorescence was measured at 440 nm and 485 nm. For plate standardization a gain of 45 was used for all measurements. Every peptide was tested in quadruplicates at two different days (n= 8). Blanks were included on every plate (n=8). As a blank control the peptide stock solution was replaced with DMSO. Data is always presented as AVG ± SD.

3.5 Virus Neutralization Assay

All infection assays were kindly performed by the laboratory of professor Dr. Schmidt (Institute of Virology and Hygiene, University of Regensburg, Regensburg). The SEAP Assays were performed by Anette Rohrhofer and Johanna Schaubächer. Luciferase Assays were done by Tamara Rügamer. Peptides were diluted from 10 mM stock solutions. Toxicity of organic solvents (DMSO and DMF) was assessed for the used concentration range and no effect was observed. Data is always presented as AVG ± SE. Blanks were always included and subtracted for the signal for data analysis. The internal positive control was the virus without peptide incubated with the same protocol and used as reference value for percentage inhibition calculation.

3.5.1 Cell Culture and Cell Lines

3.5.1.1 CEMx174-SEAP and CEMx174 M7R5 cells

For the virus neutralization two different cell lines were used. The first cell line, the CEMx174-SEAP cells were developed and kindly provided by Means and Desrosiers (Harvard Medical School, Boston, MA). These human cells were created by fusion of the T-cell line CEM and B-cell line 721.174. Both receptors, CD4 and CXCR4 are expressed on the cell surface. Quantification of infection rate with HI virus is enabled through a luminescence detection of the cell secretion of embryonic alkaline phosphatase (SEAP) which this cell line was genetically modified for. The gene is regulated *via* HIV-ltr-promoter. HIV-1 infection of the cells results in a higher secretion rate of SEAP. The cell line CEMx174M7R5 was developed by Dr. Landau (New York University Medical School, NY) and kindly provided by Dr. Korn (Institute of Virology FAU Erlangen-Nurnberg). On the cell surface CD4, CXCR4 and CCR5 are expressed. Quantification of infection rate with HIV-1 is enabled through a photometric detection of the Luciferase activity. The cell line has genes for green fluorescent protein (GFP) and Luciferase. The genes are regulated *via* HIV-ltr-promoter. HIV-1 infection of the cells results in a higher expression rate of these genes.

3.5.1.2 Cell Culture

CEMx174-SEAP and CEMx174 M7R5 cells were cultivated in RPMI 1640-medium with additives at 37 °C and 5 % CO2. In general, the cells were split twice a week and transferred to a new medium flask. CEMx174-SEAP cells were selected with G418 (100 mg/mL) every four weeks. CEMx174 M7R5 cells were permanently selected with G418 (100 mg/mL), Puromycin (5 mg/mL) and Hygromycin (50 mg/mL). Aliquots of 10-15 million cells were stored in 10 % DMSO and 90 % FBS at -80 °C. Only cells that were tested negatively for mycoplasma (with a mycoplasma detection kit) after defrosting were employed. Freshly thawed aliquots were centrifuged with fresh medium twice. The supernatant was removed after each step. Finally, the cells were resuspended and transferred to a new flask for cultivation.

3.5.1.3 HIV-1 Virus Strain Panel

For the virus cultivation 2 million infected cells in 1 mL RPMI 1640 medium with additives were incubated in 15 mL Falcon tubes at 37 °C and 5 % CO_2 with slight shaking every 20 min. The batch was transferred to a 25 cm^2 flask, filled up to 10 mL with RPMI 1640 medium with additives and incubated at 37 °C and 5 % CO_2. The culture was split and new medium was added based on the cytopathic effect. The cytopathic effect was evaluated with a light microscope every day. The culture was harvested at the peak of replication. After centrifugation in 15 mL falcon tube the supernatant was filtrated (0.22 µm pore size) and aliquoted. The virus stock aliquots were stored at -80 °C. The infection rate of the virus stocks was determined by titration. The titration was performed in 96 well cell culture plates and 25000 cells/well in RPMI 1640-medium with additives. Cells were infected with five different volumes virus stock (100 µL, 30 µL, 10 µl, 3 µl and 1 µL) yielding a final well volume of 200 µL. Negative controls without virus stocks were always included. The plate was incubated at 37 °C and 5 % CO_2 for four days. After the incubation time the plate was photometrical detected by luminescence measurement. The virus isolates used in this virus panel are listed in Table 20.

Table 20: Overview of the used virus stocks in the HIV-1 strain panel.

HIV-1-subtype	Abbreviation	Isolate	GenBank Accession Number	Coreceptor-Tropism	Source
A	A1	92UG029	AY713407	X4[1)]	NIH AIDS, Reagent Program*
A	A2	00KE_KER2018	AF457052	Dual[1)]	
CRF02_AG01	AG01	CM.0005BBY	AY371123	R5[1)]	
CRF02_AG02	AG02	01CM.0008BBY	AY371124	R5[1)]	
B	B2	90TH_BK132	AY173951	X4[1)]	
B	SF33	SF33	AY352275	X4[1)]	Jay Levy, UCSF, San Francisco
B	MN	MN	M17449	X4[1)]	NIH AIDS, Reagent Program*
B	HXBc2	HxB2	K03455	X4[1)]	
B	NL4-3	NL4-3	U26942	X4[1)]	Frank Kirchhoff, Ulm
B	BAL	93BR093	AY713409	R5[1)]	NIH AIDS, Reagent Program*
C	C1	92_BR025	U52953	R5[1)]	
C	C2	99_ET14	AY255825	R5[1)]	
D	D1	92UG021	U27399	X4[1)]	
D	D3	92UG024	U08805	X4[1)]	
F	F1	93BR029	AF005495	R5[1)]	
F	F2	93BR020	AF005494	Dual[1)]	
G	G ARP	RU570	U08368	R5[1)]	
G/H		VI525	U09665	R5[1)]	
group O	O BCF06	O BCF06	U24568	X4[2)]	
group O	O BCF011	O BCF011	U24565	X4[2)]	

*These viruses were kindly provided by Dr. Klaus Korn (Institute of Virology, FAU Erlangen-Nurnberg)

[1)] The coreceptor tropism was determined by geno2pheno [103]. Additionally, the isolates were characterized by their capability to induce syncytia in umbilical cord lymphocytes or CEMx174 M7R5 cells. Non-inducing isolates were classified as R5 viruses and inducing isolates as X4 or dual viruses.

[2)] The coreceptor tropism was only determined by geno2pheno [103].

3.5.2 Long Term Cultivation

A long-term incubation experiment was performed with PEGI peptides. For this experiment CEMx174-SEAP cells and virus isolate NL4-3 were used. The same protocol for virus cultivation was performed as described in section 3.5.1.3. Additionally, PEGI peptides were added in increasing doses. The cultivation was started in the presence of 1 µM peptide per well. The culture was split and new medium was added based on the cytopathic effect. The cytopathic effect was evaluated with a light microscope every day. If no cytopathic effect was detectable the culture was transferred together with infected cells to a new culture flask. Each time the cultivated virus was transferred the dose of peptide was increased. The final peptide concentration of the viral selection cycles was 60 µM. Finally, after all cultivation cycles the supernatant was filtered, aliquoted and stored at -80 °C. The obtained virus stocks were tested for its resistance against PEGI peptides. The virus genome was sequenced by GeneArt company (ThermoFisher, Regensburg) and analyzed with Segman™.

3.5.3 SEAP Assay

A chemiluminescent immunoassay was established to test the virus neutralization activity. The peptides of interest were diluted 1/83.3 from 10 mM stock DMSO (2.4 µL stock + 197.6 µL RPMI 1640 medium with additives). A serial two-fold dilution was performed. 10 µL diluted virus stock was added to each well and incubated at 37 °C and 5 % CO_2 for 1 h. The virus stocks were diluted depending on the virus infection which was determined previously (c.f. 3.5.1.3). Usually a 1/10 or 1/20 dilution with medium was performed. After the preincubation peptide and virus isolate were pipetted to a 96 well plate with 100 µL medium containing 25 000 seeded CEMx174-SEAP cells per well. The final volume per well was 200 µL. The cells were incubated at 37 °C and 5 % CO_2 for three days. Positive controls missing the peptide and blanks missing peptide and virus were included on each plate. Peptide or virus were replaced by the same volume of RPMI 1640-medium with additives for plate controls. After three days of incubation the chemiluminescence was measured using the Tropix Phospha-Light™ system. 12 µL supernatant per well was added to 12 µL Lysis reagent (Tropix dilution buffer) in a white 96 well Nunc™ F96 Microwell™ plate. The plate was incubated to deactivate alkaline phosphatases at 65 °C for 30 min. This was followed by adding 22 µL Phospha-Light™ Assay Buffer, incubating

at room temperature for five minutes and adding 22 µL Phospha-Light™ Reaction Buffer, incubating protected from light at room temperature for 20 min. Directly after the last incubation step the chemiluminescence was measured. Every peptide of interest was tested in triplicates.

3.5.4 Luciferase Assay

A second chemiluminescence immunoassay was established to test peptides of interest with R5 and X4 virus isolates. The peptides of interest were diluted 1/83.3 from 10 mM stock DMSO (2.4 µL stock + 197.6 µL RPMI 1640 medium with additives). A serial two-fold dilution was performed. Diluted virus stock was added to each well and incubated at 37 °C and 5 % CO_2 for 1 h. The virus stocks were diluted depending on the virus infection which was previously determined (c.f. 3.5.1.3). After the preincubation peptide and virus isolate were given to a 96 well plate with 100 µL medium containing 25 000 seeded CEMx174 M7R5 cells per well. The final volume per well was 200 µL. The cells were incubated at 37 °C and 5 % CO_2 for four days. Positive controls missing the peptide and blanks missing the peptide and virus were included on each plate. Peptide or virus was replaced by the same volume of RPMI 1640-medium with additives for plate controls. After four days of incubation the luciferase activity was measured. 40 µL resuspended cells per well and 40 µL ONE-Glo™ Luciferase Assay substrate were added to a white 96 well Nunc™ F96 Microwell™ plate and incubated for three minutes. Directly after the last incubation the Luciferase activity was measured by photometric quantification. All peptides were tested in triplicates.

3.6 Data Analysis

The following equations were used for data analysis in this work:

Average (AVG)	$AVG = \bar{x} = \frac{1}{n}\sum_{n=1}^{n} x_i$
Median	$\bar{x} = x\,\frac{n+1}{2}$
Standard deviation of the mean (SD)	$SD = \Delta x = \sqrt{\frac{\sum_{n=1}^{n}(\bar{x}-x)^2}{n(n-1)}}$
Standard error (SE)	$SE = \frac{SD}{\sqrt{n}}$
Four-parameter logistic regression	$y = y_0 + \frac{a}{1+(\frac{x}{x_0})^b}$
IC_{50} determination using four-parameter logistic regression	$IC_{50} = x_0 \sqrt[b]{\frac{a}{50 - y_0} - 1}$
IC_{30} determination based on IC_{50} **	$IC_{30} = \left(\frac{30}{100-30}^{\sqrt{hill\ slope}}\right) \times IC_{50}$
Percentage Inhibition [%]	$Inhibition\ \% = \left(1 - \frac{OD_{sample} - OD_{Blank}}{OD_{100\%} - OD_{Blank}}\right) \times 100$
Normalization of sample OD signal to a reference peptide [OD_{rel}]	$OD_{rel} = \frac{OD_{sample}}{OD_{reference}} \times 100$
Fold over X (FOX) *	$FOX = \frac{\%\ inhibition_{sample}}{\%\ inhibition_x}$
relative OD (OD_{rel})	$OD_{rel} = \frac{OD_{sample}}{OD_{reference}}$

*The fold over X (FOX) analysis was performed by Dr. Santos (TPIMS, Florida). X represents the all X sample tested for each replicate. The analysis was performed for each replicate individually and presented as median FOX per tested sample concentration.

** calculations were performed by Dr. Santos (TPIMS, Florida)

4 Results

4.1 Investigation of PEGI Peptide Activity

4.1.1 PEGIs Inhibitory Activity is based on a Prerequisite Peptide Sequence

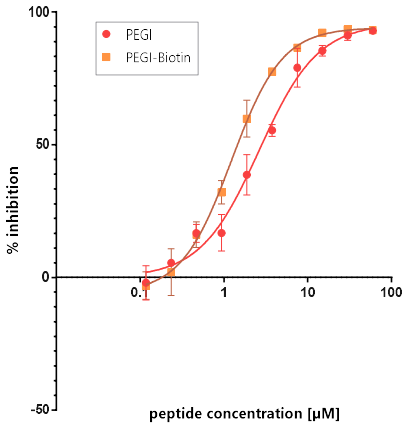
Peptides are a versatile tool to explore binding modes and sites for getting further insight into mechanism of HIV-1 entry inhibition. As a starting lead structure, the most active inhibitory peptide identified by Reil laboratory was selected (Figure 11) [22, 85]. This peptide was derived from the E2 protein of HPgV-1 virus (c.f. 1.3.3). Special characteristics of the lead peptide sequence are the hydrophobic core motif containing valine and tryptophan, three cysteines and only one charged amino acid (aspartic acid) close to the N-terminus. The aim was to investigate the main sequence characteristics causing HIV-1 inhibitory activity.

$_{45}$LCDCPNGPWVWVPAFCQAVG$_{64}$

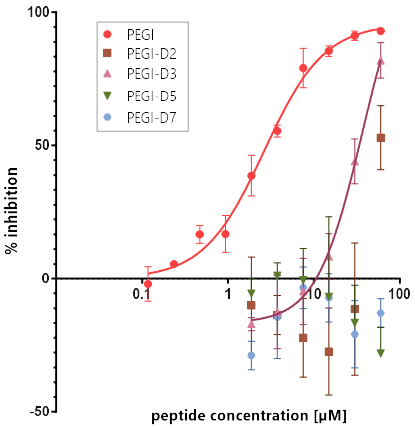
Figure 11: Lead peptide sequence of PEGI peptides (E2 protein numbering, Iowa isolate Genbank Accession Number: AF121950).

All peptides derived from this lead structure will be named and numbered as PEGI peptides. As a starting point a range of PEGI peptide variants were synthesized. The variation consisted of N-terminal biotinylation, including a N-terminal introduction of spacer AA, C-terminal functionalization and replacement of all three cysteines with serine (Figure 12 C). The introduction of biotin enabled the testing of peptide binding to monomeric recombinant viral proteins $gp120_{HXBc2}$ and $gp41_{HXBc2}$ respectively in streptavidin-based binding assays (c.f. 4.1.7).

A



B



C

% PEGI activity	
-	inactive
+	< 50 %
+ +	50-100 %
+ + +	> 100 %

		HIV-1 inhibition		Thioflavin activity
Name	Sequence	IC_{50} [µM]	% PEGI activity	n-fold gain per blank
PEGI	**Ac-LCDCPNGPWVWVPAFCQAVG-OH**	**2.8**	**lead peptide**	1.2
PEGI-D2	Bio-XLCDCPNGPWVWVPAFCQAVG-NH_2	>60	-	8.8
PEGI-D7	Ac-LSDSPNGPWVWVPAFSQAVG-NH_2	>60	-	5.4
PEGI-D5	Ac-LSDSPNGPWVWVPAFSQAVG-OH	>60	-	0.1
PEGI-D3	Ac-LCDCPNGPWVWVPAFCQAVG-NH_2	≈38	+	0.9
PEGI-Biotin	Bio-XLCDCPNGPWVWVPAFCQAVG-OH	1.5	+ + +	7.1

Figure 12: HIV-1_{NL4-3} inhibition by PEGI peptides, A: Inhibition curves of PEGI, PEGI-Biotin B: Inhibition curves of PEGI, PEGI-D2, -D3, -D5, -D7, tested in SEAP Assay C: Summary table of peptide sequence, HIV-1 inhibition and Thioflavin activity for PEGI, PEGI-Biotin, PEGI-D2, -D3, -D5 and -D7.

The SEAP assay is a HIV-1 neutralization assay that allows to test compounds or antibodies for their inhibitory activity against HIV-1 infection. It can be used to test X4-tropic HIV-1 strains. The isolate NL4-3 was the standard reference for comparison of PEGI peptide variants in HIV-1 neutralization assay. All shown HIV-1 neutralization assay experiments were performed in the laboratory of professor Dr. Schmidt (Institute of Virology and Hygiene, University of

Regensburg, Regensburg). The initial results showed that C-terminal functionalization, as well as replacement of the cysteines influenced the inhibitory activity. PEGI-D5 and -D7 (triple serine exchange variants) were completely inactive (IC_{50} > 60 µM). PEGI and PEGI-Biotin (C-terminal acid variants) were comparably active with a potent inhibitory activity in the lower micromolar range (IC_{50} = 1.5/2.8 µM). N-terminal biotinylation and additional spacer AA (PEGI-Biotin) did not influence the inhibitory activity. Consequently, the comparison of the biotinylated and non-biotinylated PEGI variants inhibitory activity was acceptable. The C-terminal amide instead reduced the inhibitory activity 20-fold (PEGI-D3) or led to complete inactivity (PEGI-D2). The obtained results confirmed the already published E2 protein-derived peptide activity seen in other HIV-1 neutralization assays. The inhibitory activity of PEGI was comparable to previously published data and the inactivity of the triple serine exchange variants (PEGI-D5 and -D7) was reproduced [22, 85]. In addition, new insight into the structural determinants for inhibitory activity could be gained. The C-terminal modification dramatically influenced the inhibitory activity of the tested PEGI peptide variants. The testing revealed that a C-terminal acid is crucial for the inhibitory activity of PEGI peptides. Moreover, a set of PEGI peptides was tested in Thioflavin assay. The hydrophobic character of the PEGI sequence, the tryptophan and valine core in particular, indicated a peptide prone to aggregation. The Thioflavin assay is a dye-based assay used for determining the aggregation potential of peptides or proteins (c.f. 3.4). A high signal increase correlates with the aggregation potential of the tested compound [100-102]. PEGI peptide variants showed different low signal gains, compared to the assay noise level, but no correlation of elevated signal in Thioflavin activity and inhibitory activity in HIV-1 neutralization assay could be deduced. PEGI and PEGI-Biotin varied in Thioflavin activity while the inhibitory activity of both peptides was comparable. Hence, the Thioflavin activity of PEGI variants was determined and an aggregation driven inhibitory activity of PEGI could be excluded.

The unexpected finding about the determining influence of the C-terminal acid for PEGI peptide variants activity required further investigation. The free acid at the C-terminus was a prerequisite for inhibitory activity, but the protein of origin (E2 protein of HPgV-1) has no natural occurring charged amino acid at this specific sequential position. To address this contradiction, extended variants and a charged side chain variant of PEGI were synthesized and tested. The 20mer peptide sequence of PEGI has a C-terminal adjacent aspartic acid within a three AA distance in the E2 protein sequence (Figure 13). Extended variants with C-terminally

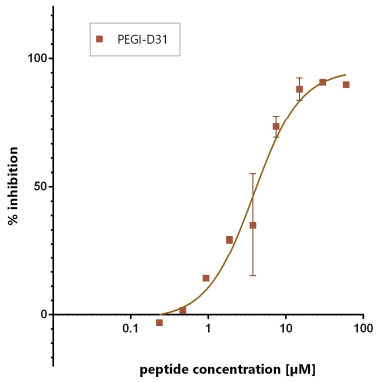
added amino acids (PEGI-D67 and PEGI-D68) were synthesized, which were corresponding to position 45-70 of the E2 protein sequence.

$_{40}$GNVTL**LCDCPNGPWVWVPAFCQAVG**WG**D**PIT$_{70}$

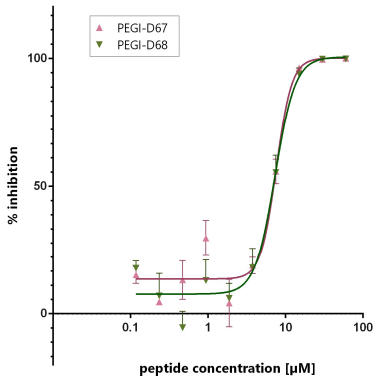
Figure 13: N-terminal section of E2-protein sequence (Iowa isolate, Genbank Accession Number: AF121950), bold black: PEGI lead sequence, bold blue: neighboring charged AA.

For PEGI-D32 the glycine at position 64 was replaced with aspartic acid and the C-terminus was amidated. This peptide showed inhibitory activity with an IC_{50} of 3.9 µM (Figure 14 A and C). The inhibitory activity was thus maintained by moving the negative charge from C-terminus to the side chain of the C-terminal AA. This indicates that its specific position is less important but its presence in the peptide sequence is mandatory. The hypothesis is confirmed by the inhibitory activity of PEGI-D67 (two AA extended variant) and -D68 (five AA extended variant). Both C-terminally extended variants of PEGI (PEGI-D67 and -D68) inhibited with IC_{50} values around 7 µM (Figure 14 B and C).

A



B



C

% PEGI activity	
-	< 10%
+	10- 50 %
+ +	50-100 %
+ + +	> 100 %

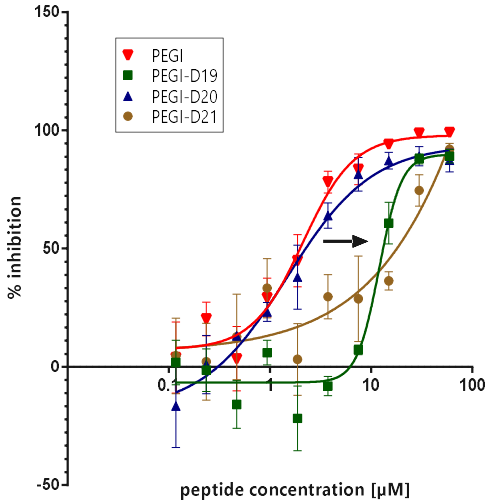
		HIV-1 inhibition	
Name	Sequence	IC_{50} [µM]	% PEGI activity
PEGI-D31	Ac-LCDCPNGPWVWVPAFCQAVD-NH_2	3.9	+ +
PEGI-D67	Bio-XLCDCPNGPWVWVPAFCQAVGWGD-NH_2	7.6	+ +
PEGI-D68	Bio-XLCDCPNGPWVWVPAFCQAVGWGDPIT-NH_2	7.3	+ +

Figure 14: HIV-1 inhibition by PEGI peptides (extended and truncated variants)
A: Inhibition curve of PEGI-D31, tested in SEAP Assay. B: Inhibition curves of PEGI-D67 and-D68, tested in SEAP Assay C: Summary table of peptide sequence, HIV-1 inhibition of PEGI-D31, -D67 and -D68.

A standard method for analyzing the contribution of the specific AA position of bioactive peptides is the random permutation of the sequence order (referred to as scrambling). Therefore, all AA were rearranged except fixed cysteines position and C-terminal negative charge (PEGI-D19). Additionally, a random peptide sequence with the same net charge and same partition coefficient (logP) (PEGI-D21) and a variant with additional fixed hydrophobic core motif (PEGI-D20) were synthesized and tested. The fixed cysteines and the fixed hydrophobic core positions were chosen due to being special characteristics of the lead sequence. The first scrambled variant PEGI-D19 (fixed cysteine position scrambled variant) was able to inhibit in HIV-1 neutralization assay, but with six-fold decrease compared to PEGI.

Apparently, the contribution of more defined AA positions than only the cysteines are needed for reaching the full inhibitory potency. PEGI-D20 (fixed cysteine and hydrophobic core position scrambled variant) exposed the importance of the position of the valine tryptophan core within the peptide sequence by maintaining the inhibitory activity of PEGI (Figure 15). The hydrophobic core was identified as an important contributing motif for PEGI activity. The random sequence control peptide (PEGI-D21) was remarkably less active with an IC_{50} value in the higher micromolar range (Figure 15). The distinct difference between PEGI and PEGI-D21 (random sequence same logP control peptide) is a strong indicator of a sequence specific inhibition mechanism of PEGI. The scrambled peptide variants confirmed the importance of the cysteine and revealed the role of the valine tryptophan core. The following structurally fundamental prerequisites for inhibitory activity could be determined for PEGI: a negative charge which can be located C-terminally or as a functionalization of the C-terminal amino acids, the three cysteines and the hydrophobic core motif of the PEGI lead peptide sequence.

A



B

% PEGI activity	
-	inactive
+	< 50 %
+ +	50-100 %
+ + +	> 100 %

		PEGI sequence	HIV-1 inhibition	
Name	Sequence	controls	IC_{50} [µM]	% PEGI activity
PEGI	**Ac-LCDCPNGPWVWVPAFCQAVG-OH**	**lead peptide**	**2.1**	
PEGI-D19	Bio-XDCLCAAWVPQVPGPNCFGWV-OH	fixed C position	12.0	+ +
PEGI-D20	Bio-XGCACDPQPFWVWVGLCPVAN-OH	fixed C + WVWV position	1.5	+ +
PEGI-D21	Bio-XIYEGCPGIAFYFWQAMNNLG-OH	random same charge + logP	≈ 25	+

Figure 15: HIV-1_{NL4-3} inhibition by PEGI peptides (scrambled variants), A: Inhibition curves of PEGI, PEGI-D19, -D20 and -D21, tested in SEAP Assay B: Summary table of peptide sequence information and HIV-1 inhibition of PEGI, PEGI-D19, -D20 and -D21.

4.1.2 Cysteine Residues play a Crucial Role for PEGI's Inhibitory Activity

Proteinogenic amino acids offer a variety of chemical functionalized side chains. Cysteine including its selenium variants bear a redox active moiety in the side chain. The methyl thioether moiety of the second prominent sulfur-containing amino acid, methionine, can be oxidized to the corresponding sulfoxide. The main cysteine reaction under physiological conditions is the oxidation to cystine *via* inter- or intramolecular disulfide formation. The standard potential (E_0) of -227 mV at pH 7.0 in aqueous solution of the cysteine-cystine redox system is described in the literature [104]. The standard potential can be considered only as an orientation benchmark for peptide or protein sequence incorporated cysteines, because neighboring amino acids or secondary structures are always influencing the redox potential [105].The disulfide bridges can be formed intra- or intermolecularly and are both described for peptides. For example, the peptide hormone insulin has an intramolecular bridge in one chain and two chains are also connected *via* a disulfide [106].The PEGI lead peptide sequence contains three cysteines which allow multiple possibilities of inter- and intramolecular disulfide bridges (Figure 16). Even combinations up to oligomers could be formed in theory. The PEGI variant, in which all cysteines were replaced with serine, was already tested and was found not to be active in HIV-1 neutralization assay (cf. 4.1.1). Therefore, PEGI peptide variants, in which one or two cysteines were replaced with serine, were synthesized and tested to explore the role of the cysteine oxidation state for the inhibitory activity.

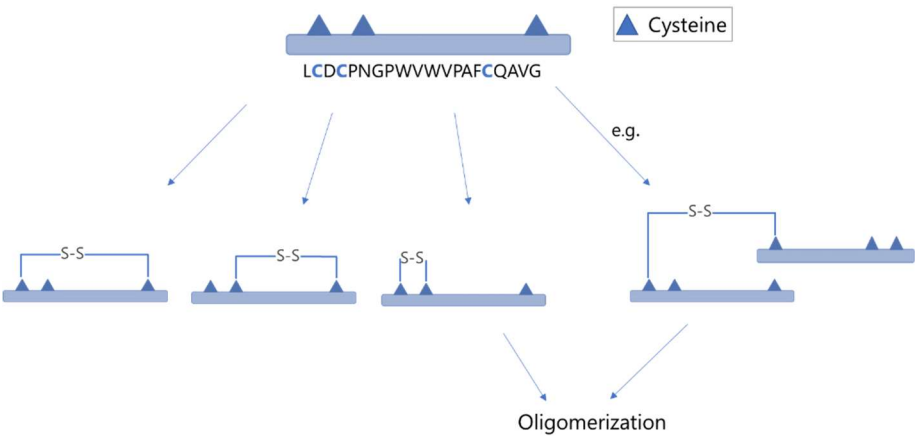


Figure 16: Schematic showing possible inter- and intramolecular disulfide bridges of a peptide with three cysteines.

Reducing the number of cysteines down to two allowed the controlled oxidation, which yielded into defined intramolecular disulfide oxidation of the two-cysteine containing peptides PEGI-D14, PEGI-D15 and PEGI-D16 (single serine exchange variants) (c.f. 3.2.2.5). PEGI peptide variants with only one cysteine were tested with a reduced thiol moiety. To limit uncontrolled oxidation processes, stock solutions were directly prepared before the first HIV-1 neutralization assay was started. The cysteine positions within the PEGI sequence are highlighted in Figure 17.

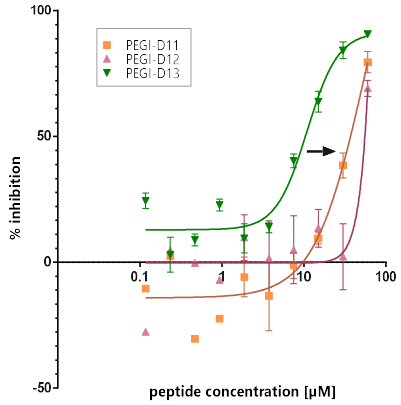
$_{45}$L**C**$_{46}$D**C**$_{48}$PNGPWVWVPAF**C**$_{60}$QAVG$_{64}$

Figure 17: Lead peptide sequence PEGI peptide, cysteines highlighted in blue.

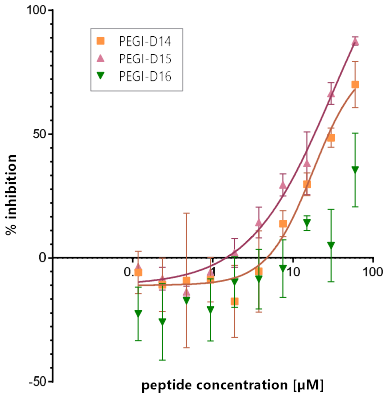
The importance of the cysteine positions was already demonstrated for the inhibitory activity of the scrambled PEGI variants. PEGI single cysteine variants with a reduced thiol moiety at position Cys_{46} and Cys_{48} (PEGI-D11 and -D12) showed a 30-fold decrease of activity compared to PEGI in HIV-1 neutralization assay, whereas inhibitory activity was almost maintained with a reduced thiol moiety at position Cys_{60} (PEGI-D13) (Figure 18 A and C). A reduced thiol moiety, non-bridged respectively, at position 60 appeared to be required for inhibitory activity. In contrast to maintained activity with only one cysteine at position 60, PEGI variants, in which one cysteine was replaced with serine (PEGI-D14, -D15 and -D16), were less active with a decrease from 10-fold to inactivity (Figure 18 B and C).

Peptides with disulfide bridges which include Cys_{60} (PEGI-D14 and -D15) lost 10 to 20-fold inhibitory activity compared to PEGI. PEGI-D16 (Cys_{60} replaced with serine) was almost inactive, but for detailed evaluation more aspects needed to be considered. The analysis of disulfide bridge formation by *Ellmans* Test and LC-MS verified that no intramolecular disulfide bridge but a dimerization occurred during oxidation of PEGI-D16 (c.f. 3.2.2.5). Even though vicinal or semivicinal disulfides can be rarely found in natural occurring proteins, the stable formation is highly sequence dependent and could not be achieved with the PEGI peptide sequence [107]. The inactivity of the PEGI-D16 in HIV-1 neutralization assay could imply that PEGIs activity is not caused by dimers connected via intermolecular disulfides and supported the results of the single cysteine variants (PEGI-D11-13) that a reduced cysteine at position 60 is required for inhibitory activity.

A



B



C

% PEGI activity	
-	< 10%
+	10- 50 %
+ +	50-100 %
+ + +	> 100 %

		Oxidation state	HIV-1 inhibition	
Name	Sequence	Cysteines	IC_{50} [µM]	% PEGI activity
PEGI-D11	Bio-XLCDSPNGPWVWVPAFSQAVG-OH	reduced	≈ 55	-
PEGI-D12	Bio-XLSDCPNGPWVWVPAFSQAVG-OH	reduced	≈ 60	-
PEGI-D13	Bio-XLSDSPNGPWVWVPAFCQAVG-OH	reduced	10.9	+ +
PEGI-D14	Ac-LCDSPNGPWVWVPAFCQAVG-OH	intramolecular disulfide	≈ 18	+
PEGI-D15	Ac-LSDCPNGPWVWVPAFCQAVG-OH	intramolecular disulfide	≈ 31	+
PEGI-D16	Ac-LCDCPNGPWVWVPAFSQAVG-OH	intermolecular disulfide *	> 60	-

*oxidation resulted in quantitative dimerization

Figure 18: HIV-1_{NL4-3} inhibition by PEGI peptides (single and double serine exchange variants) A: Inhibition curves of PEGI-D11, -D12 and -D13, tested in SEAP Assay B: Inhibition curves of PEGI-D14, -D15 and D16, tested in SEAP Assay C: Summary table of oxidation state cysteines and HIV-1 inhibition of PEGI-D11, -D12, -D13. -D14, -D15 and -D16.

The serine exchange variants provided new understanding about the role of the cysteines for PEGIs inhibitory activity. Further investigation of the functionality was still necessary. Therefore, a non-redox active and redox active surrogate approach were pursued. Methionine, a proteinogenic amino acid, with a thioether bearing side chain was used for creating redox-active cysteine exchange variants of the PEGI lead sequence. Methionine can be oxidized to its sulfoxide under physiological conditions. Consequently, no cyclization can occur. Variants containing one cysteine to methionine exchange were synthesized and tested for their

inhibitory activity. The cysteines were oxidized under formation of stable disulfide bridges, except for the variant with maintained Cys_{46}, Cys_{48} and Cys_{60} replaced with methionine (PEGI-D45). Previous experiments showed that no intramolecular disulfide bridge could be introduced and that is why the reduced PEGI-D45 was tested in HIV-1 neutralization assay. The aim was to test monomeric PEGI variants because the dimerized serine exchange variant (PEGI-D16) demonstrated that intermolecular linkage yielded in inactivity. Methionine exchange variant PEGI-D43 and PEGI-D44, both bearing an intramolecular disulfide bridge, were inactive in HIV-1 neutralization assay. The methionine exchange with reduced cysteines (PEGI-D45) was the only one that showed inhibitory activity with an IC_{50} value in the micromolar range. The triple methionine exchange variant (PEGI-D46) was also inactive (Table 21). These results supported the hypothesis that reduced thiol moieties are an essential aspect of PEGI activity and showed that methionine was a non-suitable redox surrogate for cysteine in the case of PEGI variants.

Non-redox active intramolecular cyclization was introduced into the PEGI lead sequence *via* lactam bridges. Lactam bridge variants connecting position 60 with position 46 or 48 with varying ring size were successfully synthesized. A lactam bridge between position 46 or 48 could not be synthetically introduced. No inhibitory activity was detectable for any of the three tested lactam variants (PEGI-D33, PEGI-D47 and PEGI-D48) in HIV-1 neutralization assay. Replacing disulfides with constrained lactam bridges did not yield in active PEGI variants. Not even the reduced thiol moiety at position 46 or 48 could maintain inhibitory activity (Table 22). Different conclusions could be drawn from the lactam bridge variant inactivity. It could imply that more than one reduced thiol within the sequence is needed for full activity or that cysteine at a particular position is mandatory. The cysteine at position 60 was always replaced for the introduction of a stable lactam bridge into the PEGI sequence. Moreover, it could be an indication that a flexible non-constrained peptide sequence is the active molecular species and thus inter- or intramolecular bridges are not a main requirement for its activity.

Table 21: Summary of PEGI methionine exchange variants: Sequences, oxidation states and HIV-1 inhibition of PEGI-D43, -D44, -D45 ans-D46.

% PEGI activity	
-	< 10%
+	10- 50 %
+ +	50-100 %
+ + +	> 100 %

		Oxidation state	Oxidation state	HIV-1 inhibition	
Name	Sequence	Cysteines	Methionine	IC_{50} [µM]	% PEGI activity
PEGI-D43	Bio-XLMDCPNGPWVWVPAFCQAVG-OH	intramolecular disulfide	reduced	> 60	-
PEGI-D44	Bio-XLCDMPNGPWVWVPAFCQAVG-OH	intramolecular disulfide	reduced	> 60*	-
PEGI-D45	Bio-XLCDCPNGPWVWVPAFMQAVG-OH	reduced	reduced	4.2	+ +
PEGI-D46	Bio-XLMDMPNGPWVWVPAFMQAVG-OH	-	reduced	≈ 47	-

*n=1

Table 22: Summary of PEGI lactam bridge variants: Sequences, oxidation states and HIV-1 inhibition of PEGI-D33, -D47 and -D48.

% PEGI activity	
-	< 10%
+	10- 50 %
+ +	50-100 %
+ + +	> 100 %

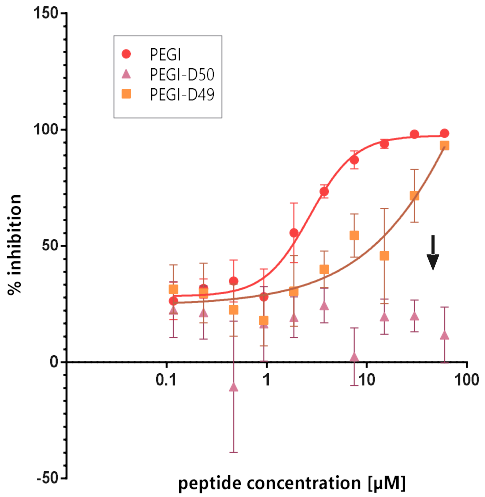
		Oxidation state	HIV-1 inhibition	
Name	Sequence	Cysteines	IC_{50} [µM]	% PEGI activity
PEGI-D47	Bio-X{KDCPNGPWVWVPAFE}QAVG-OH	reduced	> 60	-
PEGI-D48	Bio-XLCD{KPNGPWVWVPAFE}QAVG-OH	reduced	> 60	-
PEGI-D33	Ac-L-{Dap-DCPNGPWVWVPAFE}QAVG-OH	reduced	> 60	-

{ } lactam bridge

It can be noted that all variants of PEGI with defined intramolecular bridging resulted in a decrease of inhibitory activity and the correlation between activity and reduced thiol moieties was confirmed with most of the exchange variants, except the lactam bridged PEGI variants. Further details of the thiol-activity-correlation were determined using an approach with truncated variants. N-terminal truncated variants, PEGI-D49 and -D50, only bearing two cysteines were synthesized and tested in HIV-1 neutralization assay. This allowed the parallel testing of the same sequence with different defined cysteine oxidation states (reduced thiol, intramolecular disulfide). Due to the elimination of one cysteine by truncation, defined disulfide bridges could be formed. Even though the inhibitory activity of PEGI-D49 (reduced thiols) was decreased compared to PEGI, a concentration dependent inhibition was observed. The oxidized

variant PEGI-D50 (intramolecular disulfide bridge) was completely inactive. Reduced cysteines were crucial for inhibitory activity (Figure 19). This finding strongly supported the hypothesis that at least one reduced thiol is a prerequisite for PEGI activity in HIV-1 neutralization assay.

A



B

% PEGI activity	
-	< 10%
+	10- 50 %
+ +	50-100 %
+ + +	> 100 %

		Oxidation state	HIV-1 inhibition	
Name	Sequence	Cysteines	IC_{50} [µM]	% PEGI activity
PEGI	**Ac-LCDCPNGPWVWVPAFCQAVG-OH**	**lead peptide**	**2.7**	
PEGI-D49	Bio-XDCPNGPWVWVPAFCQD-NH2	reduced	≈ 20	
PEGI-D50	Bio-XD[CPNGPWVWVPAFC]QD-NH2	oxidized	> 60	-

[ ] intramolecular disulfide bridge

Figure 19: HIV-1_{NL4-3} Inhibition by PEGI variants with different oxidation states. A: Inhibition curves of PEGI, PEGI-D49 and -D50, tested in SEAP Assay. B: Summary table of oxidation state cysteine and HIV-1 inhibition of PEGI-D49 and -D50.

The role of the solvent for peptides containing oxidation sensitive amino acids like cysteine or methionine had to be carefully evaluated. DMSO was used for dissolving PEGI peptides. This solvent is suited for dissolving hydrophobic peptides, but another property can be problematic. Dimethyl sulfoxide is a mild oxidant, which is also used for cysteine oxidation [108]. That is why several precautionary measures were taken to prevent uncontrolled cysteine oxidation. The

peptides were stored in powder form at 4 °C and dissolved directly before testing in HIV-1 neutralization assay. The inhibitory activity of PEGI stock solution dissolved in another solvent was tested. DMF, which is a polar aprotic solvent without any redox activity was used in this case. The use of a different stock solution revealed an unexpected activity dependency. Inhibition curves with an IC_{50} in the lower micromolar range was only achieved with PEGI diluted from DMSO stock solutions. PEGI diluted from 10 mM DMF stock solution was only active at the highest tested concentration of 60 µM (Figure 20). Both stock solutions were further investigated. After short term storage in DMF PEGI stock solution contained visible precipitates. The *Ellmans* test, a sensitive method for free thiol detection, was conducted for both stock solutions after seven days of storage at 4 °C. A positive *Ellmans* test indicating free thiols were only detectable for DMSO stock solution. It can be hypothesized that the insoluble precipitates in DMF stock solution were oligomers formed by undirected cysteine oxidation. It can be concluded that oxidation in DMF stock solution caused inactivity by probable intermolecular linkage up to oligomerization. PEGI dissolved in DMSO remained active under the tested storage conditions. These results supported that not all three cysteine of PEGI were completely oxidized when dissolved in DMSO. The number of free thiols per peptide remained undetermined. The stock solvent plays a crucial role for PEGI activity and the investigation confirmed the correlation of reduced thiols and PEGI activity in HIV-1 neutralization assay.

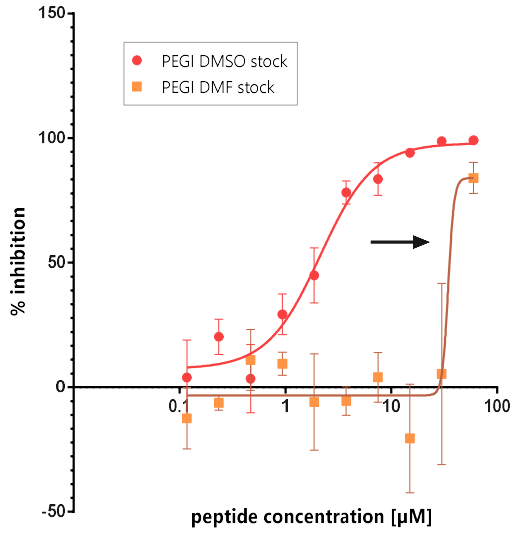
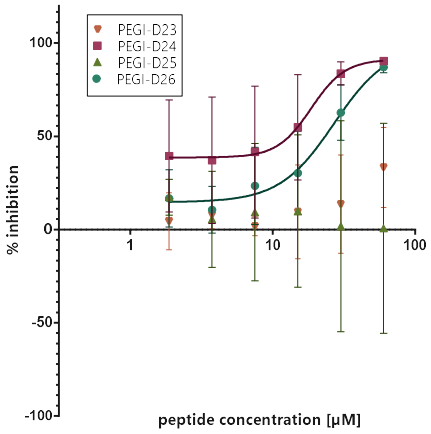


Figure 20: HIV-1_{NL4-3} inhibition by PEGI diluted from DMSO and DMF stock solutions, inhibition curves of PEGI from 10 mM stock solution in DMSO and DMF, tested in SEAP Assay.

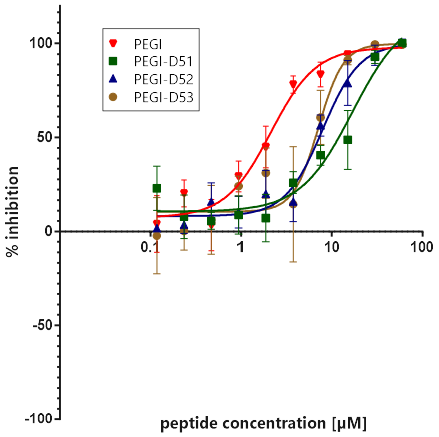
4.1.3 PEGI Key Structure Analysis by Truncation and Exchange Variants Including D-Amino Acids, Alanine and Homocysteine

The initial determination of functional prerequisites for PEGIs inhibitory activity with scrambled peptide variants emphasized the importance of the hydrophobic core of the sequence (c.f. 4.1.1). Further structure activity analysis was performed. The functional role of hydrophobic parts of antiviral peptides had been described for HIV-1 entry inhibitors [109-111]. The consecutive replacement of all amino acids with alanine (alanine scan), is a robust method to determine the contribution of the specific side chain to the overall sequential activity and was therefore applied to PEGI hydrophobic core. The alanine scan of the motif was synthesized and tested in HIV-1 neutralization assay. As expected, the inhibitory activity was distinctly affected. PEGI-D23 and PEGI-D25 ($W_{53/55}$ to A exchange variants) were inactive and the activity of PEGI-D24 and PEGI-D26 ($V_{54/56}$ to A exchange variants) was 10- to 20-fold decreased (Figure 21 A). Losing the aromatic heterocycle side chain of tryptophan, abolished the inhibitory activity completely. Tryptophan is known to contribute to peptide-protein interactions *via* π-stacking on the molecular level [7]. Starting from these results, the effect of changing the stereochemistry of selected amino acids was investigated by introduction D-configurated building blocks into the PEGI lead sequence. The inhibitory activity range was maintained for PEGI-D51 and -D52 (L-Trp to D-Trp variants). Introducing D-tryptophan at position 55 (PEGI-D51) resulted in a 9-fold lower IC_{50} value compared to PEGI but only a 4-fold lower one for position 53 (PEGI-D52) (Figure 21 B). Additionally, a variant with D-aspartic acid at position 47 was tested (PEGI-D53). The position was selected because of the two adjacent cysteines. The inhibitory activity was almost maintained with PEGI-D53 (Figure 21 B). The stereochemistry of the selected AA appeared to play no major role for inhibitory activity but L-tryptophan was favorable at position 55. The alanine scan and D-amino acid exchange variant data is summarized in Figure 21 C.

A



B



C

% PEGI activity	
-	< 10%
+	10- 50 %
+ +	50-100 %
+ + +	> 100 %

		Exchange	HIV-1 inhibition	
Name	Sequence	Cotnrols	IC_{50} [µM]	% PEGI activity
PEGI	**Ac-LCDCPNGPWVWVPAFCQAVG-OH**	**lead peptide**	**2.1**	
PEGI-D23	Ac-LCDCPNGP**A**VWVPAFCQAVG-OH	W → A	> 60	-
PEGI-D24	Ac-LCDCPNGPW**A**WVPAFCQAVG-OH	V → A	18.6	+
PEGI-D25	Ac-LCDCPNGPWV**A**VPAFCQAVG-OH	W → A	>60	-
PEGI-D26	Ac-LCDCPNGPWVW**A**PAFCQAVG-OH	V → A	26.9	+
PEGI-D51	Bio-XLCDCPNGPWV**w**VPAFCQAVG-OH	W → w	16.9	+ +
PEGI-D52	Bio-XLCDCPNGP**w**VWVPAFCQAVG-OH	W → w	7.9	+ +
PEGI-D53	Bio-XLC**d**CPNGPWVWVPAFCQAVG-OH	D → d	7.1	+ +

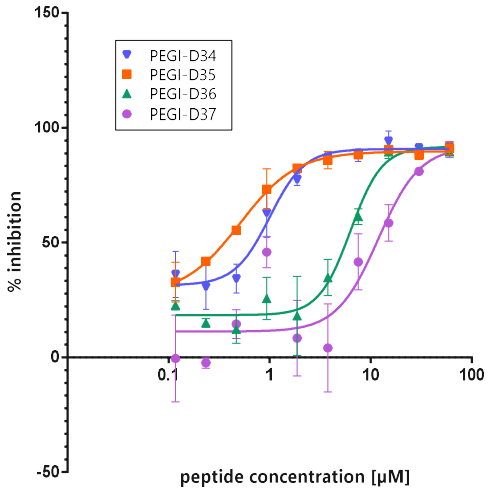
lower case letter = D-amino acid

Figure 21: HIV-1_{NL4-3} inhibition by PEGI peptides (single AA exchange variants) A: Inhibition curves of PEGI-D23, -D24, -D25 and -D26, tested in SEAP Assay. B: Inhibition curves of PEGI, PEGI-D51, -D52 and-D53, tested in SEAP Assay C: Summary table of peptide sequences, exchange variant and HIV-1 inhibition of PEGI-D23, -D24, -D25 and -D26 and PEGI-D51, -D52 and-D53.

Another well-established method, beside alanine scanning, was applied to determine the minimal sequence length needed for full inhibitory activity. It has been demonstrated that the C-terminal negative charge can be switched to the side chain (c.f. 4.1.1). On this basis, truncation variants with a C-terminal aspartic acid amide were synthesized and tested. The results showed the methodical efficacy because the minimal length of the active PEGI sequence

was successfully determined. The inhibitory activity of PEGI was not only maintained but slightly increased with omission of two C-terminal amino acids. C-terminal truncation of more than two AA or any N-terminal truncation resulted in activity decrease (Figure 22 A).

A



B

% PEGI activity	
-	< 10%
+	10- 50 %
+ +	50-100 %
+ + +	> 100 %

		Truncation	HIV-1 inhibition	
Name	Sequence	Site	IC_{50} [µM]	% PEGI activity
PEGI-D34	Bio-XLCDCPNGPWVWVPAFCQAD-NH_2	C-terminal -1	1.0	+ + +
PEGI-D35	Bio-XLCDCPNGPWVWVPAFCQD-NH_2	C-terminal -2	0.5	+ + +
PEGI-D36	Bio-XLCDCPNGPWVWVPAFCD-NH_2	C-terminal -3	6	+ +
PEGI-D37	Bio-XCDCPNGPWVWVPAFCQAVG-OH	N-terminal -1	11.9	+ +

Figure 22: HIV-1_{NL4-3} by PEGI peptides (C- and N-terminal truncation variants) A: Inhibition curves of PEGI-D34, -D35, -D36 and -D37, tested in SEAP Assay B: Summary table of peptide sequences, binding and HIV-1 inhibition of PEGI-D34, -D35, -D36 and -D37.

PEGI sequence truncation allowed the determination of the minimal sequence length required for maximal inhibitory activity. Moreover, the inhibitory activity could be increased 5-fold compared to PEGI. The 17mer sequence of PEGI-D35 (C-terminal truncated variant) was a substantial improvement of the neutralizing capacity. The investigation of the structural

functionality of the PEGI lead sequence resulted in the discovery of a more potent variant (Figure 23). The new discovered more potent PEGI sequence variant is suited for further testing as a new lead compound.

$_{45}$L**C**$_{46}$D**C**$_{48}$PNGPWVWVPAF**C**$_{60}$QD$_{62}$-NH$_2$

Figure 23: PEGI-D35 sequence, minimal length required for maximal inhibitory activity, cysteine positions are highlighted in blue.

The potent PEGI-D35 sequence was further used for cysteine position analysis in addition to previous studies described in chapter 4.1.2. In this case the focus was oriented towards stereochemistry and side chain length of the thiol moiety. Exchange variants with a single introduced D-cysteine were synthesized and tested in HIV-1 neutralization assay. The introduction of a single D-amino acid did not result in inactive peptides but the position of the D-cysteine influenced notably the inhibitory activity. PEGI-D61, D-cysteine at position 60, had a low inhibitory activity with high variation compared to PEGI. PEGI-D62 (Cys_{46} D-Cys exchange variant) and -D63 (Cys_{48} D-Cys exchange variant) showed a more consistent concentration response dependency with IC_{50} values in the lower micromolar range (Figure 25 A). In general, D-configured cysteine variants were less active but especially the C-terminal L-configured-cysteine at position 60 appeared to be essential for maintaining inhibitory activity of PEGI-D35.

A non-proteinogenic cysteine homologue was used to explore the influence of the distance between backbone to thiol moiety on the inhibitory activity. Therefore, variants with homocysteine, whose sidechain is extended by one methylene spacer, were synthesized and tested (Figure 24).

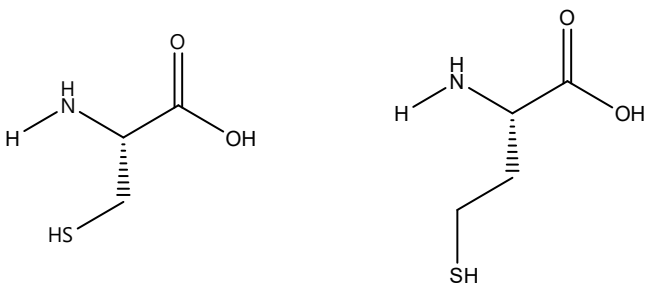
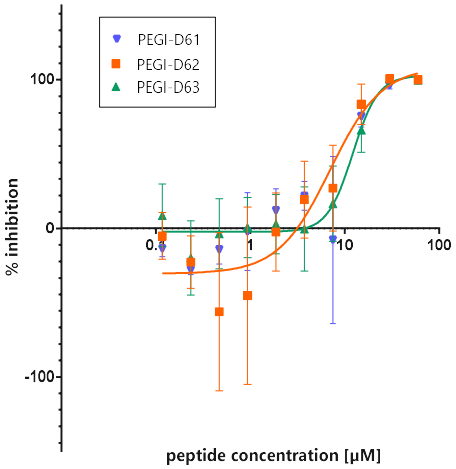


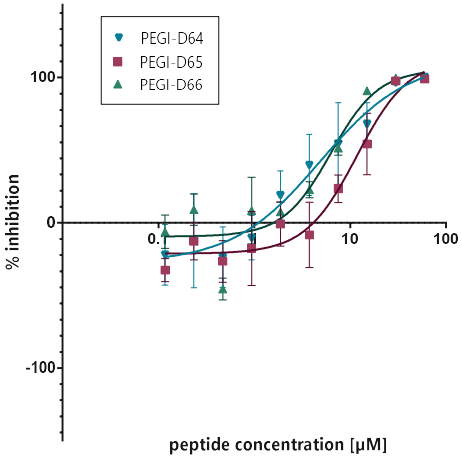
Figure 24: Structural formula of cysteine and homocysteine.

In contrast to the D-AA cysteine exchange variants, the implementation of homocysteine showed no position preference. The IC_{50} values of PEGI-D64, -D65 and -D66 (homocysteine exchange variants) ranged from 4.9 to 11.8 µM and thus lower inhibitory activity was achieved compared to PEGI-D35 (Figure 25 B). This emphasized the importance of the three thiol containing sidechains with their specific chain length. Both PEGI-D35 modification approaches confirmed the special role of the cysteines (Figure 25 C).

A



B



C

		HIV-1 inhibition	
Name	Sequence	IC_{50} [µM]	% PEGI activity
PEGI-D61	Bio-XLCDCPNGPWVWVPAF**c**QD-NH_2	active*	+
PEGI-D62	Bio-XL**c**DCPNGPWVWVPAFCQD-NH_2	7.0	+ +
PEGI-D63	Bio-XLCD**c**PNGPWVWVPAFCQD-NH_2	12.8	+ +
PEGI-D64	Bio-XL-**Hcy**-DCPNGPWVWVPAFCQD-NH_2	4.9	+ +
PEGI-D65	Bio-XLCD-**Hcy**-PNGPWVWVPAFCQD-NH_2	11.8	+ +
PEGI-D66	Bio-XLCDCPNGPWVWVPAF-**Hcy**-CQD-NH_2	6.3	+ +

% PEGI activity	
-	< 10%
+	10- 50 %
+ +	50-100 %
+ + +	> 100 %

*curve could not be fitted

lower case letter = D-amino acid, Hcy = Homocysteine

Figure 25: HIV-1_{NL4-3}. Inhibition by PEGI-D35 D-Cysteine and Homocysteine exchange variants. A: Inhibition curves of PEGI-D61, -D62 and -D63, tested in SEAP Assay. B: Inhibition curves of PEGI-D64, -D65 and -D66, tested in SEAP Assay. C: Summary table of peptide sequences and HIV-1 inhibition of PEGI-D61, -D62, -D63, -D64, -D65 and -D66.

4.1.4 PEGI Resistant HIV-1_{NL4-3} showed Peptide Specific Mutation Pattern

Long-term incubation experiments were performed to get a better understanding for the underlying mechanism of HIV-1 inhibition by PEGI. The cultivation of HIV-1 isolates in the presence of inhibitory compounds enabled the development of compound resistant viruses. Changes in viral genes of interest of these viruses can be analyzed by sequencing. In this case three aliquots of HIV-1_{NL4-3} were incubated over 18 or 23 passages with either increasing concentrations up to 60 µM of PEGI, PEGI-D25 or no peptide (c.f. 3.5.2). PEGI-D25 (W_{55} to A exchange variants) was selected as non-inhibitory peptide control (control 1) (cf. 4.1.3). The second incubation control was performed without any additional peptide. Instead of a peptide solution buffer was used in this case (control 2). The number of passages of the experiment depended on the viral susceptibility to inhibition by PEGI. The used concentration of peptides was limited due to cytotoxic effects of DMSO when peptides were applied in concentrations higher than 60 µM. After the long-term cultivation was finished the virus isolates were tested in standard HIV-1 neutralization assay for its susceptibility to inhibition by PEGI. Simultaneously, non-incubated NL4-3 (wildtype) was tested as an additional control. In the following obtained virus isolates will be called: virus passage 18, passage 23, control 1 (incubation with PEGI-D25), control 2 (incubation without peptide) and wildtype (no incubation). The virus incubated in the presence of PEGI was tested at passage 18 and 23. Only virus passage 23 was resistant against PEGI after the long-term incubation procedure. Virus control 1 was still effectively inhibited by PEGI. The virus passage 18, control 2 and wildtype were also susceptible to inhibition by PEGI (Figure 26). The distinct change of susceptibility to inhibition by PEGI evolved after passage 18 of cultivation in the presence of PEGI.

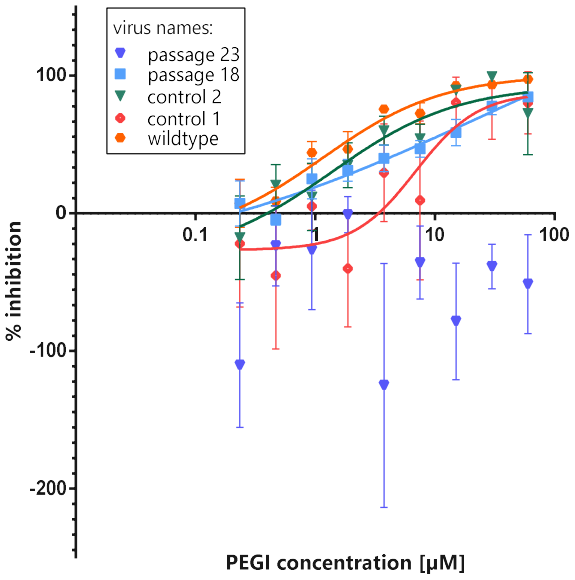


Figure 26: HIV-1 inhibition of the long-term incubated viruses by PEGI. Inhibition curves of virus passage 18 and 23 (cultivated in the presence of PEGI), virus control 1 (cultivated in the presence of PEGI-D25), virus control 2 (cultivated without peptide) and non-incubated reference NL4-3 (wildtype), tested in SEAP Assay.

The genes of ENV of the virus cultivated in the presence of PEGI were sequenced after passage 6, 18 and 23. After passage 6, one mutation in the N-terminal region of gp120 (K33N) was observed. The following sequencing revealed three additional mutations in gp120 after passage 18: two in the C1 region (V120E, S128N) and one in the V3 loop (A327T). The virus was still non-resistant but bearing five single amino acid mutations within the gp120 sequence at passage 18. The final mutation appeared in the V2 region (F175L) after passage 23. The virus gained resistance against PEGI peptide (60 µM). After passage 18 gp41 was also sequenced and a mutation in the transmembrane region was observed (R185K) (Table 23). The genes of the two control viruses were also sequenced. Single mutations were detected: D57G and A273T (control 1, PEGI-D25) and E349K (control 2, no peptide) (Table 23). It should be noted that the controls did not show a single mutation in similar gp120 parts compared to the virus cultivated with PEGI. Observed mutations of the virus cultivated in the presence of PEGI are further illustrated in Figure 27. Mutation K33N is located in the N-terminal part of gp120, whereas the three additional mutations (V120E, S128N, F175L) are in closer proximity in the C1 and V2 region. The crystal structure used for analysis showed the envelope protein in a prefusion trimer conformation [38]. Based on this specific crystal structure mutation site F175L is not surface

exposed (Figure 27). The mutation A327T can be found at the V3 stem but its surface exposition is diminished in comparison to the highly exposed V3 loop tip (Figure 28). The mutation at position 185 in gp41 is located in the transmembrane region which is anchored into the viral membrane and no structural data is available for this section of gp41 so far. The long-term incubation in the presence of PEGI resulted in a resistant virus and control virus isolates were still susceptible to inhibition by PEGI. Additionally, gp120 sequencing revealed five mutations which did not appear within the control isolates.

Table 23: Single amino acid mutations of long-term incubated NL4-3 in the presence of PEGI, control 1 (PEGI-D25) and control 2 (no peptide) (NL4-3 numbering).

Incubated NL4-3	Passage	Mutations gp120					Mutation gp41
PEGI	6	K33N	-	-	-	-	n.d.
	18	K33N	V120E	S128N	A327T	-	n.d.
	23	K33N	V120E	S128N	A327T	F175L	R185K
control 1	18	D57G	A273V	-	-	-	n.d.
control 2	18	E349K	-	-	-	-	n.d.

n.d. = not determined

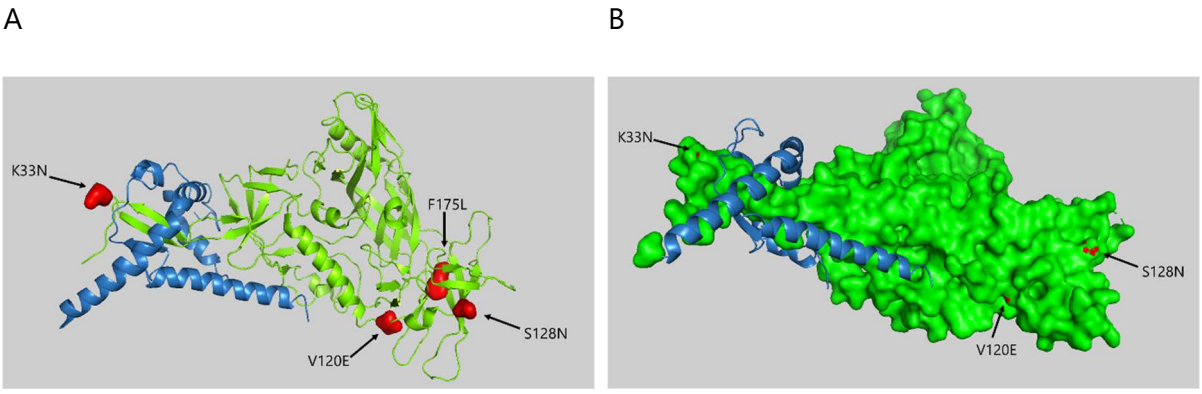


Figure 27: Crystal structure showing gp120 (green), gp41 (blue) and three mutations of virus passage 23 (red) (pdb: 4TVP) A: Structure shown as cartoon style B: Structure of g41 shown as carton style and gp120 shown with covered surface.

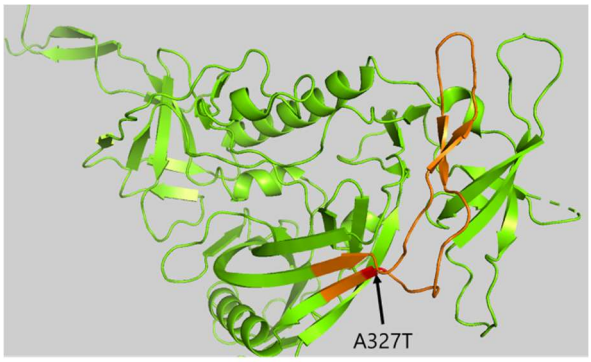


Figure 28: Crystal structure section showing gp120 (green) with highlighted V3 loop (orange) and mutation at the V3 loop base A327T (red) (pdb: 4TVP).

4.1.5 PEGI Lead Sequence is a Conserved Motif Among HPgV-1 Strains

The PEGI lead sequence represents the amino acids 45 to 64 of the E2 protein of the Iowa strain of HPgV-1 (Genbank Accession Number: AAD31765). Since the discovery of the virus, sequences of various HPgV-1 strains have become available. Using a protein blast approach, 250 available E2 protein sequences in the NCBI databank were analyzed [112, 113]. The strain analysis revealed that the PEGI lead peptide sequence is highly conserved within the available sequences and only three amino acid positions vary among analyzed strains (Figure 29).

$$_{45}\text{LCDCPNGPWVW}X_{56}\text{PA}X_{59}\text{CQA}X_{63}\text{G}_{64}$$

Figure 29: PEGI lead sequence variable positions are highlighted in blue (numbering: E2 protein Iowa strain (Genbank Accession Number: AAD31765)), variable positions were determined *via* a NCBI databank search.

Based on the sequence analysis, single and combined exchange variants of position 56, 59 and 63 were synthesized and tested in HIV-1 neutralization assay (Table 24). A Luciferase assay was used which enables the testing of X4 and R5 tropic HIV-1 isolates. The relevance will be further discussed in section 4.1.6. Especially the incorporation of amino acids with a different chemical character inside the hydrophobic core motif was of big interest (PEGI-D55, -D56, -D57

and -D58). The importance of the core motif for inhibitory activity was already shown with alanine exchange variants of PEGI (c.f. 4.1.3).

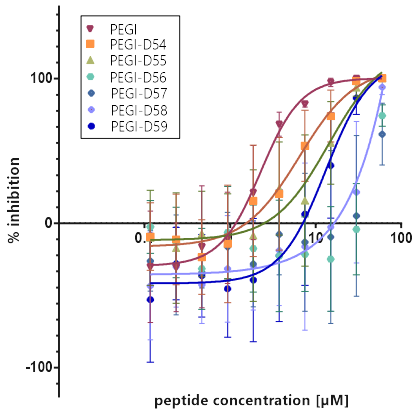
Table 24: Variable positions of E2 protein NCBI databank search. n=250 (number of sequence entries that were analyzed).

Position	Amino acid and number of ocurrence					
56	L 223	V 17	H 2	Q 3	S 1	M 2
59	F 235	L 13				
63	I 212	V36				

n= 250

Strain variants with single AA and double AA exchanges (PEGI-D54, -D55, -D56, -D57, -D58 and -D59) were tested for their inhibitory activity. All tested PEGI HPgV-1 strain variants were less active than PEGI. The inhibitory activity of PEGI as seen before in SEAP assay (IC_{50} = 2.8 µM) was successfully reproduced with the Luciferase assay system (IC_{50} = 2.3 µM). The expression of CCR5 and CXCR4 of the used cell line and the different detection system did not influence PEGI's activity. PEGI-D54 (V56L and V63I exchange variant) with an IC_{50} of 6.4 µM reached almost the potency of PEGI, followed by moderate activity of the other PEGI-D55 (V56M exchange variant) and PEGI-D59 (F59L exchange variant) (Figure 30 A). The exchange of valine with leucine or methionine at position 56 (PEGI-D54 and -D55) was tolerated whereas glutamine, serine or histidine instead of valine at position 56 (PEGI-D56, -D57 and-D58) caused a decline in inhibitory activity (Figure 30 B). PEGI and PEGI-D54 were representing the most occurring sequence variants of the E2 sequence NCBI databank search. Both peptides showed concentration dependent inhibition of HIV-1_{NL4-3} infection with higher potency compared to the other tested strain variants. Consequently, the activity of PEGI lead peptide sequence obtained from Iowa strain of HPgV-1 (Genbank Accession Number: AAD31765) was not a strain specific sequence phenomenon but a prevalent highly conserved motif. PEGI exchange variants with valine at position 56 replaced with methionine or leucine and variants with phenylalanine at position 50 replaced with leucine maintained inhibitory activity.

A



B

% PEGI activity	
-	< 10%
+	10- 50 %
+ +	50-100 %
+ + +	> 100 %

		HIV-1 inhibition	
Name	Sequence	IC_{50} [µM]	% PEGI activity
PEGI	Ac-LCDCPNGPWVWVPAFCQAVG-OH	2.3	
PEGI-D54	Bio-XLCDCPNGPWVW**L**PAFCQA**I**G-OH	6.4	+ +
PEGI-D55	Bio-XLCDCPNGPWVW**M**PAFCQAVG-OH	14.3	+ +
PEGI-D56	Bio-XLCDCPNGPWVW**Q**PAFCQAVG-OH	active*	+
PEGI-D57	Bio-XLCDCPNGPWVW**S**PAFCQAVG-OH	active*	+
PEGI-D58	Bio-XLCDCPNGPWVW**H**PAFCQAVG-OH	active*	+
PEGI-D59	Bio-XLCDCPNGPWVWVPA**L**CQAVG-OH	13.1	+ +

*curves could not be fitted

Figure 30: HIV-1_{NL4-3} inhibition by PEGI peptides (HPgV-1 strain variants) A: Inhibition curves of PEGI, PEGI-D54, -D55, -D56, -D7, -D58 and -D59, tested in Luciferase Assay B: Summary table of peptide sequences and HIV-1 inhibition of PEGI, PEGI-D54, -D55, -D56, -D57, -D58 and -D59.

4.1.6 PEGI Inhibits HIV-1 Isolates Belonging to Different Clades

The standardized protocol for testing PEGI and its derivates used a HIV-1 isolate called NL4-3. It belongs to HIV-1 clade B and uses the CXCR4 coreceptor for host cell entry. For testing the breadth of HIV-1 inhibitory activity, five additional strains belonging to clade B were tested in HIV-1 neutralization assay. In this case a Luciferase assay was used which enables the testing of X4- and R5-tropic HIV-1 isolates in the same assay system. PEGI was able to inhibit all clade B strains with IC_{50} values in the lower micromolar range in a comparable manner to the results obtained in SEAP assay with NL4-3 (c.f. 4.1.1). The R5-tropic isolate BAL was susceptible to inhibition by PEGI in a comparable manner as the tested X4-tropic isolates of clade B. Moreover, non-clade B isolates were available and subsequently PEGIs ability to inhibit HIV-1 infection was tested in a HIV-1 strain panel composed of 20 different strains (Figure 31). The tested strains from clade A, B, F, G (G/H) were susceptible to inhibition by PEGI with IC_{50} values between 0.08-4.5 µM [114]. PEGI even inhibited two HIV-1 isolates from group O with IC_{50} values in the lower micromolar range. The HIV-1 inhibitory activity of PEGI was only moderate against isolate D1 and for isolate D3 only an insufficient concentration dependency was obtained. Nevertheless, both clade D isolates were susceptible to inhibition by PEGI at the tested higher concentrations. Four strains, both CRF02-AG isolates, F1 and C2, were resistant against inhibition by PEGI in the tested concentration range. All resistant strains were R5-tropic. X4-, R5- and dual tropic strains were successfully inhibited and PEGI appeared to show a preference for inhibiting X4-tropic isolates.

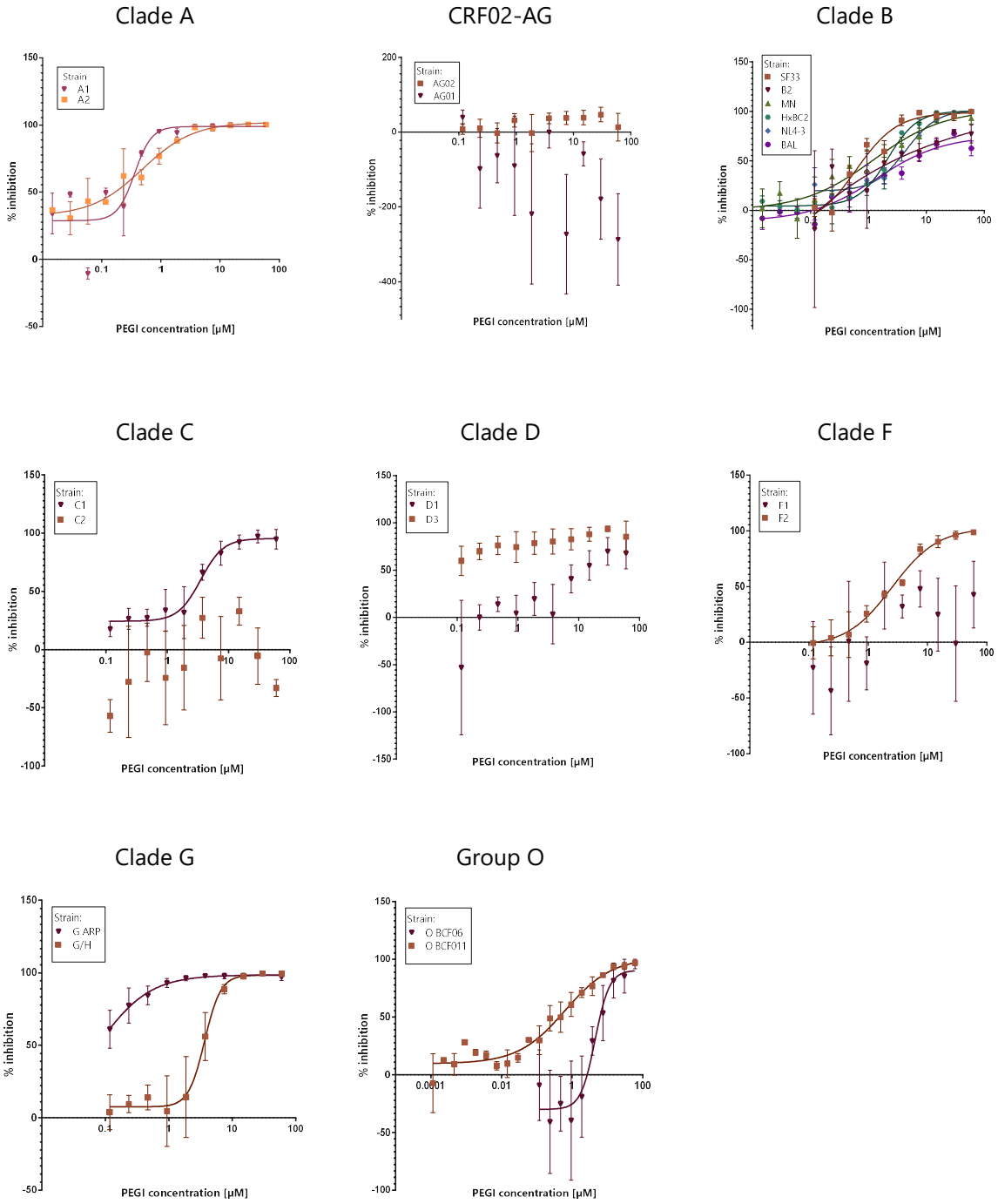


Figure 31: HIV-1 strain panel results of HIV-1 inhibition by PEGI. Inhibition curves of PEGI sorted by clade, strains belonging to a clade are presented together in a graph. Testing was performed in Luciferase assay, each concentration per isolate was tested at least in duplicates.

The HIV-1 strain panel results were further validated with an authority approved drug, called T20 (Enfuvirtide), which inhibits the HIV-1 entry to its host cell. All 20 isolates were susceptible to inhibition by T20 in nanomolar concentrations, which is in accordance with the HIV-1 inhibitory activity range described in literature [115, 116] (Chapter 7.2, Table 42). Apart from the successful validation of the HIV-1 strain panel, the specificity of the AG01, AG02, F1 and C2 resistance against inhibition by PEGI was thereby supported. The results of the HIV-1 strain panel data for PEGI are summarized in Table 25. The resistance of four isolates (AG01, AG02, F1 and C2) could imply that a common viral characteristic was provoking it. However, a coincidence as well as an additional isolate specific effect could also be plausible. In addition it could be shown that HIV-2 strains were not susceptible to inhibition by PEGI [114].

Table 25: Overview of HIV-1 inhibition results of PEGI, tested in the HIV-1 strain panel in Luciferase assay, IC_{50} values >60 µM highlighted in orange.

HIV-1 strains				PEGI
Abbreviation	Isolate	HIV-1-subtype	Tropism	IC_{50} [µM]
A1	92UG029	A	X4	0.4
A2	00KE_KER2018	A	dual	0.5
AG01	CM.0005BBY	CRF02_AG	R5	>60
AG02	01CM.0008BBY	CRF02_AG	R5	>60
B2	90TH_BK132	B	X4	I active *
SF33	SF33	B	X4	0.6
MN	MN	B	X4	1.2
HxBC2	HxB2	B	X4	2.2
NL4-3	NL4-3	B	X4	4.4
BAL	93BR093	B	R5	1.4
C1	92_BR025	C	R5	3.5
C2	99_ET14	C	R5	>60
D1	92UG021	D	X4	III active *
D3	92UG024	D	X4	I active *
F1	93BR029	F	R5	>60
F2	93BR020	F	dual	2.5
G ARP	RU570	G	R5	0.1
G/H	VI525	G/H	R5	3.7
O BCF06	O BCF06	group O	X4	4.5
O BCF011	O BCF011	group O	X4	0.6

*curves could not be fitted and HIV-1 inhibitory activity grouped by concentration with 50 % inhibition

active*	50 % inhibition [µM]
I	1-10
II	10-30
III	30-60

To get a better understanding about the relationships and sequence homology of the HIV-1 strain panel, ENV sequence alignment and phylogenetic analysis was performed. The focus was on the envelope protein because of its role for host cell entry (c.f. 1.2.2). The results of the phylogenetic analysis showed that the panel represented a broad spectrum of ENV sequences. PEGI was not only able to inhibit closely related isolates, but effective inhibition was achieved for 16 of the 20 tested isolates. The PEGI resistant isolates AG01 and AG02 are closely related but C2 and F1 are not directly neighboring the CRF branch of the phylogenetic tree. The distance of the branches represents the proportion of sequential homology of the individual isolates (Figure 32).

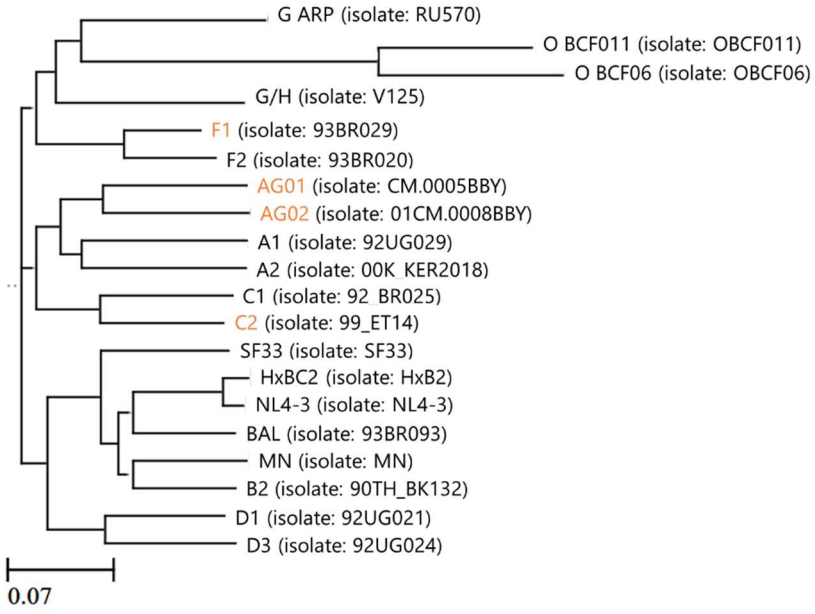


Figure 32: Phylogenetic tree of HIV-1 strain panel, created with clustal omega webtool and ETE toolkit [117-120], PEGI resistant isolates are highlighted in orange.

4.1.7 PEGI Binds to Monomeric Recombinant Envelope Proteins – a Mechanism for the HIV-1 Inhibitory Activity

The detailed investigation of the structural prerequisites as well as neutralization capacity and breadth presented a potent HIV-1 inhibitory peptide. However, the mechanism for the HIV-1 inhibitory activity remained unresolved. Therefore, a panel of PEGI variants was tested in

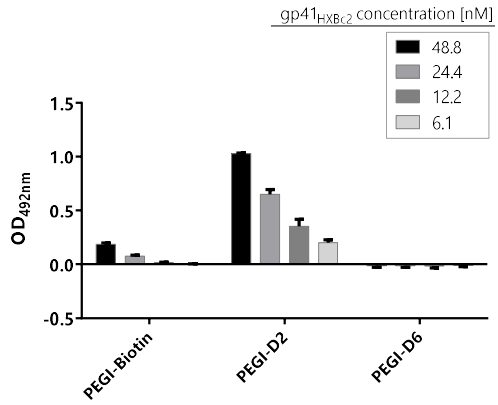
binding assays to determine the correlation between binding strength and HIV-1 inhibitory activity. Established ELISA-based binding assays using monomeric recombinant $gp41_{HXBc2}$ $gp120_{HXBc2}$ were used to study the binding of PEGI and its derivates. Gp120 and gp41 proteins from HXBc2 were selected as standard references due to their high ENV sequence homology with NL4-3 (97 %), which was used as standard reference in HIV-1 neutralization assay. The sequence homology was determined by HIV-1 isolate sequence alignment [113]. A set of biotinylated PEGI peptides that was used to study the HIV-1 inhibitory activity was also tested in streptavidin-based binding assays (Table 26).

Table 26: Biotinylated PEGI peptide sequences and characteristics.

Name	Sequence	Characteristic
PEGI-Biotin	Bio-XLCDCPNGPWVWVPAFCQAVG-OH	C-terminal acid
PEGI-D2	Bio-XLCDCPNGPWVWVPAFCQAVG-NH_2	C-terminal amide
PEGI-D6	Bio-XLSDSPNGPWVWVPAFSQAVG-NH_2	triple serine variant
PEGI-D32	Bio-XLCDCPNGPWVWVPAFCQAVD-NH_2	C-terminal exchange G → D
PEGI-67	Bio-XLCDCPNGPWVWVPAFCQAVGWGD-NH_2	C-terminal + 3
PEGI-68	Bio-XLCDCPNGPWVWVPAFCQAVGWGDPIT-NH_2	C-terminal + 5

The binding assay results showed a higher binding strength of PEGI-D2 (C-terminal amide) to both recombinant ENV proteins compared to PEGI-Biotin (C-terminal acid). For PEGI-Biotin low binding signals were observed at the highest concentration tested of $gp41_{HXBc2}$ binding assay. The concentration range was three-fold higher for $gp41_{HXBc2}$ than for $gp120_{HXBc2.}$ Whereas the C-terminal amid peptide (PEGI-D2) binding signals were similar for both recombinant proteins, the C-terminal acid peptide (PEGI-Biotin) showed a binding preference to $gp120_{HXBc2}$. The carboxy group at the C-Terminus of PEGI-Biotin appeared to influence the binding selectivity of the peptide sequence. The binding as well as the HIV-1 inhibitory activity were influenced by the C-terminal charge of the peptides (c.f. 4.1.1). Based on these results, a gp41 binding involving mechanism for HIV-1 inhibition by PEGI is not very likely. The binding strength indicated that the HIV-1 inhibition could be linked by PEGI-Biotins ability to bind gp120. PEGI-D6, in which all three cysteines were replaced with serine, was completely inactive in both binding assays (Figure 33). The inactivity of the triple serine variant (PEGI-D6) in $gp41_{HXBc2}$ and $gp120_{HXBc2}$ binding correlated with the inactivity of all tested triple serine variants in HIV-1 neutralization assay.

A



B

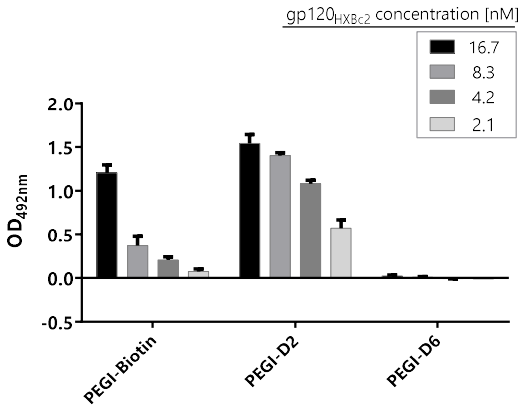
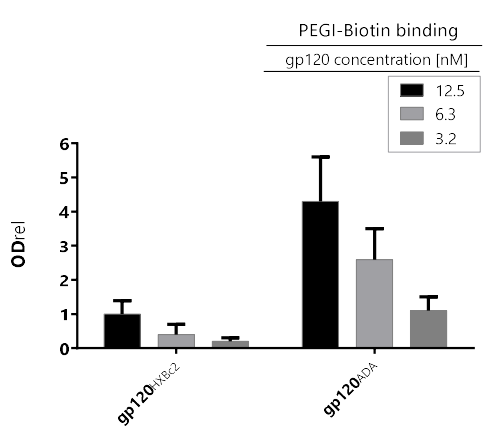


Figure 33: Binding to recombinant ENV proteins A: Binding of biotinylated PEGI peptides to $gp120_{HXBC2}$ (16.7/8.3/4.2/2.1 nM), tested in Assay Format A B: Binding of biotinylated PEGI peptides to $gp41_{HXBC2}$ (48.8/24.4/12.2/6.1 nM), tested in Assay Format A.

For further verification of PEGIs ability to bind gp120, different clade B derived recombinant gp120 proteins were tested. PEGI-Biotin showed concentration dependent binding to gp120 belonging to isolates MN and ADA. The recombinant $gp120_{HXBc2}$ was included in each experiment and used as a control. In comparison to $gp120_{HXBc2}$, higher binding strength to $gp120_{ADA}$ and similar one to $gp120_{MN}$ was observed (Figure 34). This could be an indication that a binding site of PEGI is located in a conserved region of ENV because the binding was not limited to a single strain derived recombinant gp120 protein. The binding to a R5-tropic protein implied that a tropism selective binding is not to be expected. The binding of PEGI to $gp120_{ADA}$ confirmed the HIV-1 inhibition by PEGI against several R5-tropic HIV-1 isolates tested in the HIV-1 strain panel (c.f. 4.1.6).

A



B

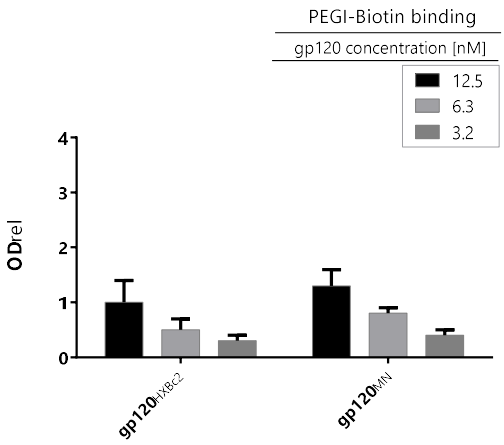


Figure 34: Binding to recombinant gp120 proteins A: Binding of PEGI-Biotin to $gp120_{HXBc2}$ and $gp120_{ADA}$ (12.5/6.3/3.2 nM), Binding Assay Format A B: Binding of PEGI-Biotin to $gp120_{HXBc2}$ and $gp120_{MN}$ (12.5/6.3/3.2 nM) tested in Binding Assay Format A, OD_{rel} was calculated with OD_{492} signal of 12.5 nM PEGI-Biotin binding to $gp120_{HXBc2}$.

In the following it was determined if the HIV-1 inhibitory activity of PEGI variants PEGI-D31 (truncated variant), PEGI-67 (two AA extended variant) and PEGI-D68 (five AA extended variant) also correlated with the ability to bind $gp120_{HXBc2}$. For the binding assay the biotinylated variant of PEGI-D31 (PEGI-D32) was used. The three PEGI variants (PEGI-D32, -D67 and -D68) bound to $gp120_{HXBc2}$ with a higher strength compared to the initial lead peptide PEGI-Biotin. (Figure 35).

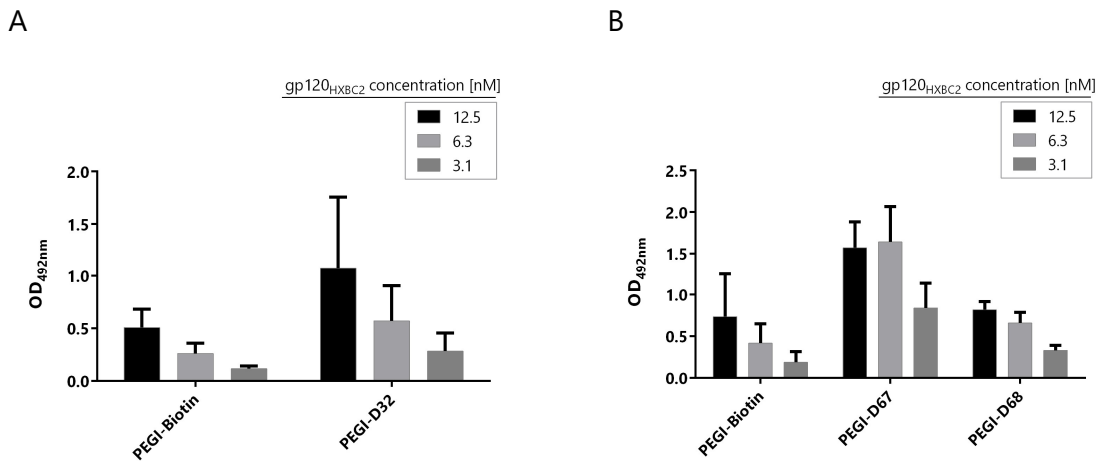


Figure 35: Binding to recombinant gp120$_{HXBc2}$ A: Binding of PEGI-Biotin and PEGI-D32 to gp120$_{HXBc2}$ B: Binding of PEGI-Biotin and PEGI-D67 and -D68 to gp120$_{HXBc2}$, tested in Binding Assay Format A.

The binding assay results correlated with HIV-1 inhibitory activity of the tested PEGI derivates (c.f. 4.1.3). Moreover, the correlation revealed that the gp120 could be the molecular target for HIV-1 inhibition by PEGI. The specific binding site yet remained unclear.

4.1.8 PEGI Inhibits Binding of HIV-1 Neutralizing Antibodies to $gp120_{HXBc2}$

The activity of PEGI was investigated on a protein binding level. It could be shown that the biotinylated triple serine variant (PEGI-D4) was inactive whereas a concentration dependent binding of PEGI-Biotin was observed (c.f. 4.1.7). The specific binding site remained unclear. Nevertheless, the binding studies showed a selectivity of the active PEGI peptides for $gp120_{HXBc2}$ compared to $gp41_{HXBc2}$. To get further insight into the PEGI binding site, competitive ELISA studies were conducted. The aim was to test the ability of PEGI to interfere with the interaction of HIV-1 antibodies with their antigen proteins. Based on the known epitopes of the antibodies, inhibition indicates a putative binding site of PEGI. A panel of ten ENV-targeting antibodies was tested. These antibodies can be classified by their epitopes into several groups. For this purpose, robust and reliable assays with individual controls for each antibody were established. The competitive assay was verified with positive and negative controls. The standard inhibition control was $gp120_{HXBc2}$ itself for all gp120-targeting antibody competitive assays and the epitope representing peptide $Disulfide\text{-}Loop_{HXBc2\text{-}ox}$ was used for $F240\text{-}gp41_{HXBc2}$ competitive assay (cf. 7.2, Figure 66 and Figure 67). In addition, PEGI-D5, the non-binding triple serine exchange variant, was always included as a negative control in the testing. Three different CD4 binding site antibodies (b12, VRC01 and 3BNC117) were tested in competitive assay. PEGI inhibited the CD4-binding site antibody-$gp120_{HXBc2}$ interaction whereas PEGI-D5 (triple serine exchange variant) was inactive (Figure 36). In particular, inhibition curves for VRC01 and 3BNC117 reached a plateau and no saturation was reached for b12 in the applied concentration range of PEGI. Due to the limited solubility of the peptides in assay buffer, higher concentrations could not be tested. PEGIs activity was limited to 60 % inhibition of the VRC01-and 3BNC117-$gp120_{HXBc2}$- interaction. That is why the inhibitory activity was compared at 5 µM concentration of PEGI (Table 27). Micromolar concentrations of PEGI accomplished half maximal inhibition of the CD4 binding site antibody-$gp120_{HXBc2}$ interactions, but complete inhibition was not achieved.

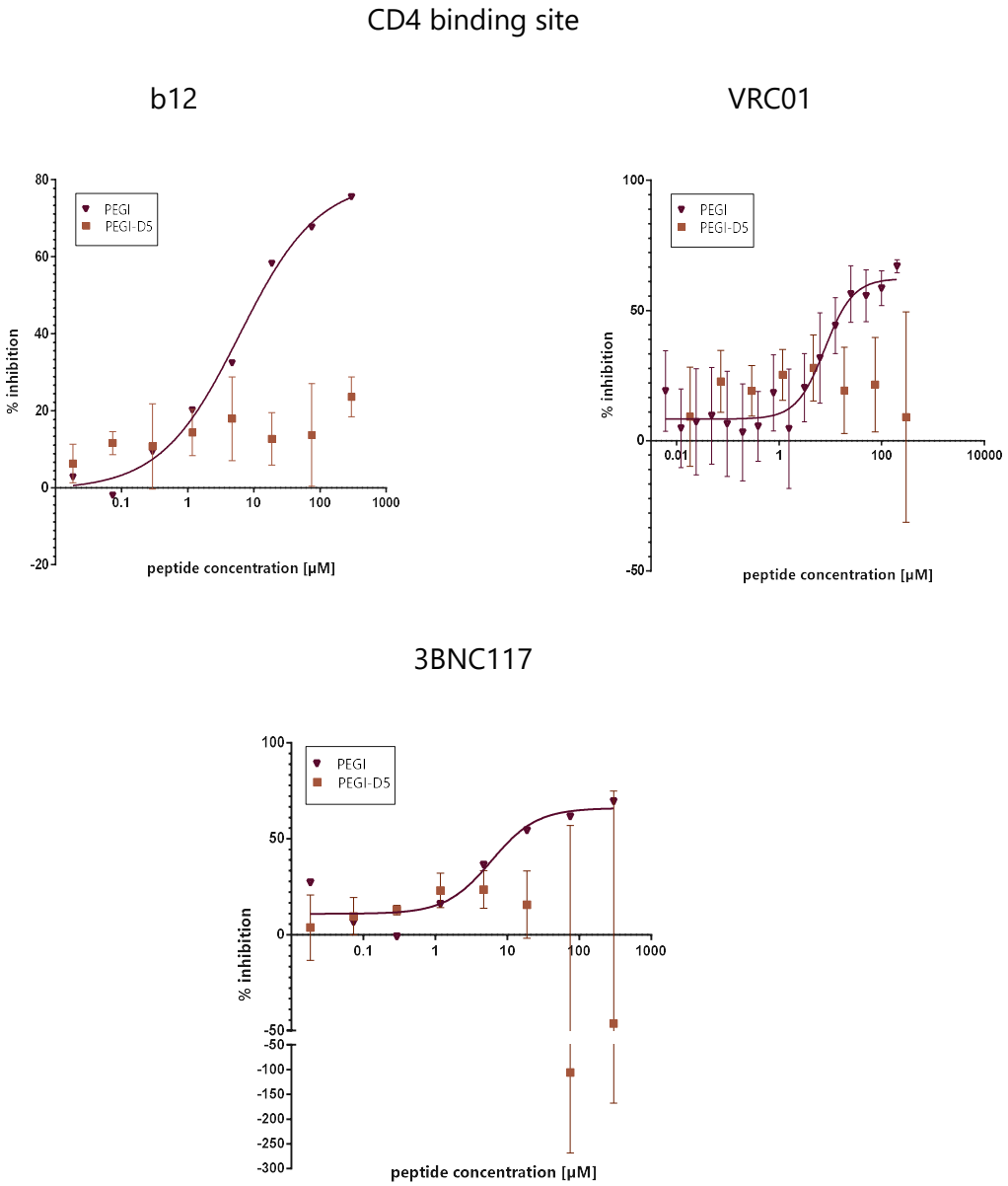


Figure 36: Inhibition of CD4-binding-site antibody binding to $gp120_{HXBc2}$ by PEGI peptides. Concentration dependent inhibition curves of PEGI and PEGI-D4, tested in Competitive Assay Format D with CD4 binding site antibody b12 (0.1 µg/mL), VRC01 (0.1 µg/mL), and 3BNC117 (0.02 µg/mL).

The V3 loop is a surface exposed region of gp120 that is targeted by several monoclonal antibodies. This variable part plays a crucial role for coreceptor selectivity of the individual virus (c.f. 1.2.2). Competitive assay systems were established for two human and one murine

monoclonal antibody targeting the V3 loop of $gp120_{HXBc2}$. The binding of $gp120_{HXBc2}$ to all three antibodies could be successfully inhibited by PEGI. PEGI-D5 (triple serine exchange variant) instead was inactive. An inhibition of approximately 70 % was achieved for mAb 447-52D and 5F7. The inhibition of F425 B4e8-$gp120_{HXBc2}$ interaction was even more effective because PEGI was able to fully inhibit the interaction with a concentration dependent curve and an IC_{50} of 2 µM. The three tested V3 loop antibodies were sensitive to inhibition by PEGI in the applied concentrations but only for F425 B4e8 100 % inhibition was reached (Figure 37). The obtained response curves indicated a considerable higher accessibility of the F425 Be48-gp120 interaction for inhibition by PEGI compared to the two other tested V3 loop antibodies.

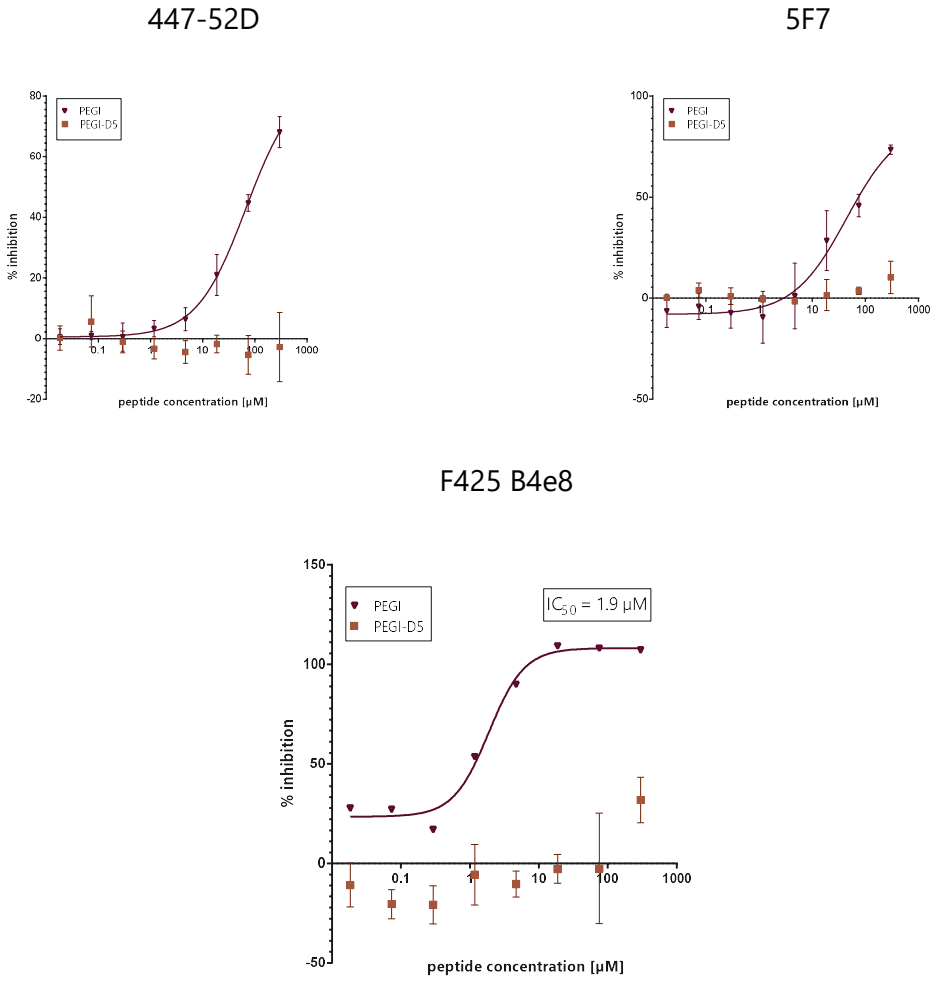


Figure 37: Inhibition of V3 loop antibody binding to $gp120_{HXBc2}$ by PEGI peptides. Concentration dependent inhibition curves of PEGI and PEGI-D5, tested in Competitive Assay Format D with 447-52D (0.01 µg/mL), 5F7 (1/200 dilution) and F425 B4e8 (1.5 µg/mL).

A glycan motif located at the V3 stem is the known epitope of the broadly HIV-1 neutralizing antibody 2G12. The ability of PEGI and PEGI-D5 to inhibit the 2G12-$gp120_{HXBc2}$ interaction was tested. Both peptides failed to inhibit the 2G12-$gp120_{HXBc2}$ interaction although the 2G12 competitive assay was verified with an inhibition control (Figure 38 and chapter 7.2, Figure 66). Hence, the resistance against inhibition by PEGI was not caused by the used assay system but rather peptide specific. Consequently, V3 loop neighboring glycans can be considered to be not a part of the putative PEGI binding site.

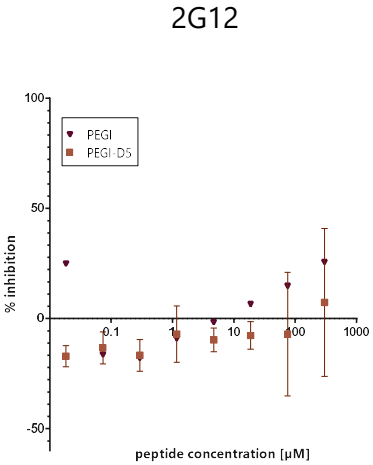


Figure 38: Inhibition of 2G12 binding to $gp120_{HXBc2}$ by PEGI peptides. Concentration dependent inhibition curves of PEGI and PEGI-D5, tested in Competitive Assay Format D with 2G12 (0.05 µg/mL).

ID6 is a murine HIV-1 non-neutralizing gp120 binding antibody with an epitope within N-terminal region. The ID6-gp120$_{HXBc2}$ interaction could not be inhibited by PEGI or PEGI-D5, although the interaction was sensitive to inhibition by the inhibition control (Figure 39 and chapter 7.2, Figure 66). Thus, the N-terminal region of gp120 being a candidate for PEGI binding site was not very likely.

ID6

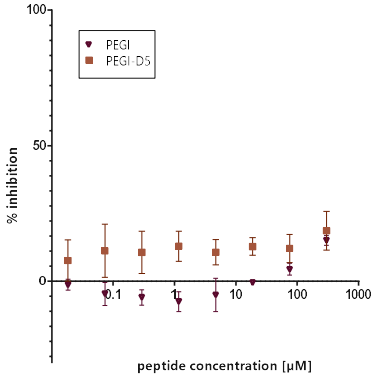


Figure 39: Inhibition of ID6 binding to gp120$_{HXBc2}$ by PEGI peptides. Concentration dependent inhibition curves of PEGI and PEGI-D5, tested in Competitive Assay Format D with ID6 (1/1000 dilution).

The coreceptor binding site of gp120 is addressed by several antibodies. As a representative mAb X5 was selected and a competitive assay was established. PEGI-D5 showed no good concentration dependent activity, but an inhibition around 30-40 % was observed when tested in higher concentrations. On the contrary PEGI showed concentration dependent activity but no inhibition over 65 % was achieved (Figure 39). For this antibody no conclusive differentiation between the two peptides was observed because both peptides showed low to moderate inhibitory activity. Compared to the other tested antibodies of the panel the inhibitory activity of PEGI in $gp120_{HXBC2}$-X5 competitive assay was low. It can be assumed that a PEGI binding site within the X5 epitope is not very likely to be the main reason for its inhibitory activity in HIV-1 neutralization assay.

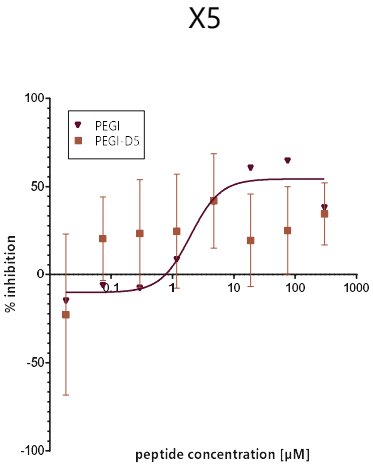


Figure 40: Inhibition of X5 binding to $gp120_{HXBc2}$ by PEGI peptides. Concentration dependent inhibition curves of PEGI and PEGI-D5, tested in Competitive Assay Format D with X5 (2 µg/mL).

Beside gp120 the second ENV sub-unit gp41 is another major antibody target. MPER targeting antibodies like 4E10 and 2F5 target a conformational fusion step called six-helix-bundle formation [40]. That is why these antibodies could not be tested with recombinant monomeric $gp41_{HXBc2}$. Nevertheless, a different highly conserved epitope of $gp41_{HXBc2}$ sequence could be addressed with the antibody F240. Unlike all the other tested antibodies, PEGI-D5 (triple serine variant) showed a concentration dependent inhibition curve with an IC_{50} of 2.7 µM. PEGI showed only 50 % inhibition in concentrations over 100 µM (Figure 40). Biotinylated variant of PEGI-D5 (PEGI-D6) showed neither binding to $gp41_{HXBc2}$ nor to F240 (c.f. 4.1.7, chapter 7.2, Figure 65 G). Based on these results, it is questionable if inhibition by PEGI-D5 is caused by a specific binding to $gp41_{HXBc2}$.

F240

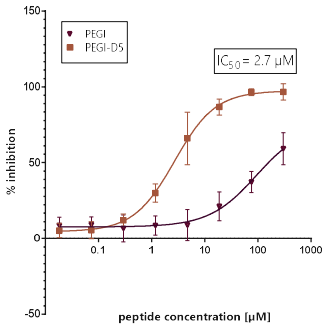


Figure 41: Inhibition of F240 binding to $gp41_{HXBc2}$ by PEGI peptides. Concentration dependent inhibition curves of PEGI and PEGI-D5, tested in Competitive Assay Format C with F240 (0.017 µg/mL).

The antibody panel comprised one gp41 and nine gp120 antibodies. The ability of PEGI and PEGI-D5 (triple serine exchange variant) to bind any of these antibodies was determined by Streptavidin-based binding assays using their biotinylated variants. In the tested concentrations no binding of biotinylated variant of PEGI (PEGI-Biotin) and PEGI-D5 (PEGI-D4) to any antibody was detected (chapter 7.2, Figure 65). Based on these results, direct interaction with the antibodies could be excluded as the reason for the observed activity. Thus, inhibition by PEGI was very likely due to its ability to bind gp120 or gp41. The competitive assay results showed that PEGI was capable of inhibiting the $gp120_{HXBc2}$ interaction of several HIV-1 neutralizing antibodies. The triple serine variant PEGI-D5 instead showed almost no concentration dependent inhibition, except for the F240-$gp41_{HXBC2}$ and X5-$gp120_{HXBc2}$

interaction. Although a number of interactions could be inhibited by PEGI, the efficacy varied considerably. The antibodies can be grouped into two main groups: sensitive and resistant against inhibition by PEGI. 2G12 and ID6 were unambiguously resistant. The sensitive group can be further divided into low (F240, X5), moderate (447-52D and 5F7) and high (3BNC117, VRC01, b12 and F425 B4e8) sensitive interactions. 100 % inhibition with an IC_{50} of 2 µM was only achieved for V3 loop antibody F425 B4e8. For the tested $gp41_{HXBc2}$ antibody F240 the peptide activity was not selective for PEGI (Table 27).

Table 27: Summary of the antibody panel: binding site, neutralizing potency and inhibition results of PEGI.

Antibody	Binding site	Broadly Neutralizers*	Inhibition by 4.7 µM PEGI
			AVG [% inhibition]
3BNC117	CD4 binding site	x	36**
VRC01	CD4 binding site	x	25**
b12	CD4 binding site	x	47
X5	Coreceptor binding site on gp120		41**
F240	Disulfide loop of gp41		9
2G12	Glycan-V3 site on gp120	x	< 5
ID6	N-Terminus of gp120		< 5
F245 B4e8	V3-Loop gp120		90
447-52D	V3-Loop gp120	(x)	6
5F7	V3-Loop gp120		1

* classification based on [29]

** curves saturated around 50-60 % inhibition

4.1.9 PEGI interacts with F425 B4e8 on the epitope level

The testing of an HIV-1 antibody set discovered the effective inhibition of F425 B4e8 $gp120_{HXBc2}$ interaction by PEGI (c.f. 4.1.8). The binding epitope of F425 B4e8 is located in the tip region of the V3 loop of gp120 [121]. For further investigation of this particular inhibition a peptide – peptide binding assay was established. The binding assay was set up as a streptavidin-based ELISA and that is why the biotinylated variants of PEGI (PEGI-Biotin) and the triple serine exchange variant PEGI-D5 (PEGI-D4) were used. The cyclic peptide $V3\text{-}Loop_{HXBc2}$ displays the exact amino acid sequence of this specific part of isolate HXBc2 (Figure 42).

Biotin-XGX-[CTRPNNNTRKRIRIQRGPGRAFVTIGKIGNMRQAHC]-NH_2

Figure 42: $V3\text{-}Loop_{HXBc2}$ peptide sequence, V3 loop section is highlighted in blue.

The binding of PEGI-Biotin and PEGI-D4 to the peptide $Fluo\text{-}V3\text{-}Loop_{HXBc2}$ was tested in binding assay. To facilitate the binding assay the Fluorescein-labeled variant was used ($Fluo\text{-}V3\text{-}Loop_{HXBc2}$). It could be demonstrated that PEGI binds to $Fluo\text{-}V3\text{-}Loop_{HXBc2}$ in nanomolar concentrations but PEGI-D4 (biotinylated triple serine exchange variant) showed no binding (Figure 43). The strong binding to $Fluo\text{-}V3\text{-}Loop_{HXBc2}$ was only detectable for the cysteine containing PEGI-Biotin and correlated with its binding to monomeric $gp120_{HXBc2}$. Biotinylated triple serine exchange variants (PEGI-D4 or -D6) instead were not active in either $V3\text{-}Loop_{\text{-}HXBc2}$ or $gp120_{HXBc2}$ binding assay (cf. 4.1.7).

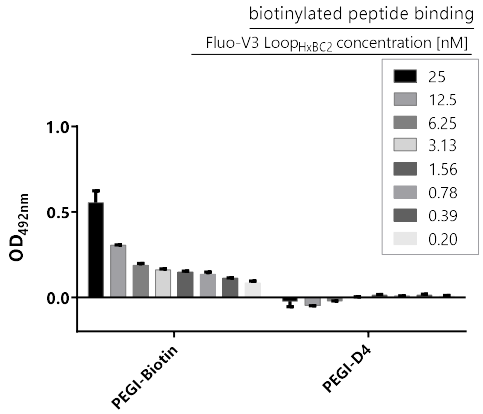


Figure 43: Binding of PEGI-Biotin and PEGI-D4 (biotinylated triple serine exchange variant) to Fluo-V3 Loop$_{HXBc2}$ peptide, tested in Binding Assay Format C in a two-fold serial dilution starting at 25 nM, n=2.

Moreover, the strong binding of PEGI to the Fluo-V3-Loop$_{HXBc2}$ peptide was the starting point to investigate if the F425 B4e8-gp120 interaction could be displayed with a peptide antibody assay. Thus, a competitive F425 B4e8-V3-Loop$_{HXBc2}$ assay was established, verified by positive control and PEGI peptides were tested (chapter 7.2, Figure 68). The F425 B4e8-V3-Loop$_{HXBc2}$ assay confirmed the inhibitory activity of PEGI. PEGI inhibited the interaction with a convincing concentration response and an IC_{50} of 3.2 µM, but PEGI-D4 was completely inactive (Figure 44). PEGI inhibited the binding of F425 B4e8 to V3-Loop$_{HXBc2}$ (IC_{50} = 3.2 µM) and monomeric gp120$_{HXBc2}$ (IC_{50} = 1.9 µM) with comparable efficacy (c.f. 4.1.7). This coherence strengthened the hypothesis that PEGIs inhibitory activity against V3 loop antibodies might be caused by its ability to bind the epitope region of mAb F425 B4e8.

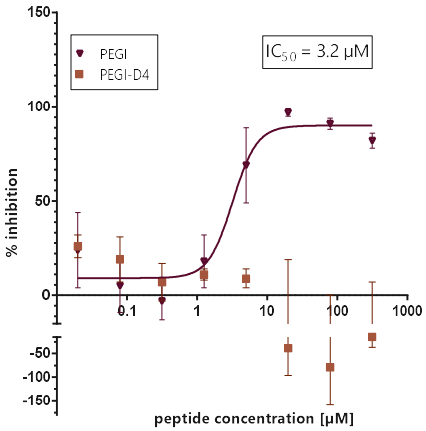


Figure 44: Inhibition of F425 B4e8-V3-Loop$_{HXBc2}$ interaction by PEGI peptides. Concentration dependent inhibition curves of PEGI and PEGI-D4, tested in Competitive Assay Format E with F425 B4e8 (4.7 nM).

Binding and competitive assays with monomeric gp120$_{HXBc2}$ are robust systems to investigate peptide protein interaction, even though the naturally occurring complex conformational changes or trimerization of ENV could not be simulated with monomeric ENV proteins. That is why the transferability of these findings to the level of HIV-1 neutralization was further investigated. Based on the fact that PEGI binds to the V3 loop of HXBc2 displayed in a peptide, the next step was to test whether the V3-Loop$_{HXBc2}$ peptide could interfere with HIV-1 inhibition by PEGI. The performed experiment is illustrated in Figure 45. 10 µM PEGI inhibited 80 % of the HIV-1 infection. This inhibition could be counteracted by the addition of V3-Loop$_{HXBc2}$ peptide, which was tested at six concentrations ranging from 0.1 to 7.5 µM. It resulted in a partly restoration of the initially inhibited HIV-1 infection. The HIV-1 inhibitory effect of PEGI was halved by the addition of 7.5 µM V3-Loop$_{HXBc2}$. Higher concentrations of V3-Loop$_{HXBc2}$ could not be tested because of its own HIV-1 inhibitory activity (chapter 7.2, Figure 69). Due to this effect the assay was only performed with an excess of PEGI compared to V3-Loop$_{HXBc2}$. It can be hypothesized that the restoration of HIV-1 infection is due to the capture of PEGI by the V3-Loop$_{HXBc2}$ and hence the binding to the viral gp120 is prevented, as well as consequently the inhibition of HIV-1 infection. The result showed that the ability of PEGI to bind the V3 loop is also present in the context of HIV-1 neutralization assays.

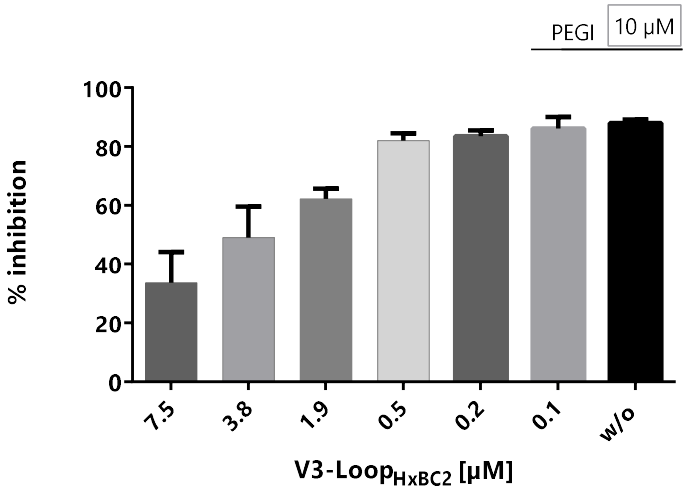


Figure 45: Inhibition of HIV-1_{NL4-3} by PEGI and partly restoration of HIV-1 infection by additionally added V3-$Loop_{HXBc2}$ peptide. Each column representing the HIV-1_{NL4-3} inhibition by PEGI without (w/o) or with V3-$Loop_{HXBc2}$ peptide (0.1/0.2/0.5/1.9/3.8/7.5 µM), tested in SEAP Assay.

4.2 Peptide Library Scan with HIV-1 Antibodies

Since the beginning of HIV-1 research, antibody-protein interactions have been intensively studied to understand the complex viral infection and to identify new targets, compounds and immunogens. The search for new inhibitory compounds targeting a specific antibody epitope or protein region involved in the HIV-1 replication cycle is an important aspect of the HIV-1 research. Usually the target structure of an antibody is referred to as antigen [122]. In the case of one important group of HIV-1 neutralizing antibodies the antigenic epitopes are located in the envelope protein (ENV) of the virus. For the inhibition of an antibody-protein interaction, two binding mechanisms are possible. The inhibitory compound, in this case a peptide, could either inhibit by binding to the antibody or the protein (Figure 46). The binding results in the shielding of the interaction of antibody and its defined epitope and subsequent inhibition.

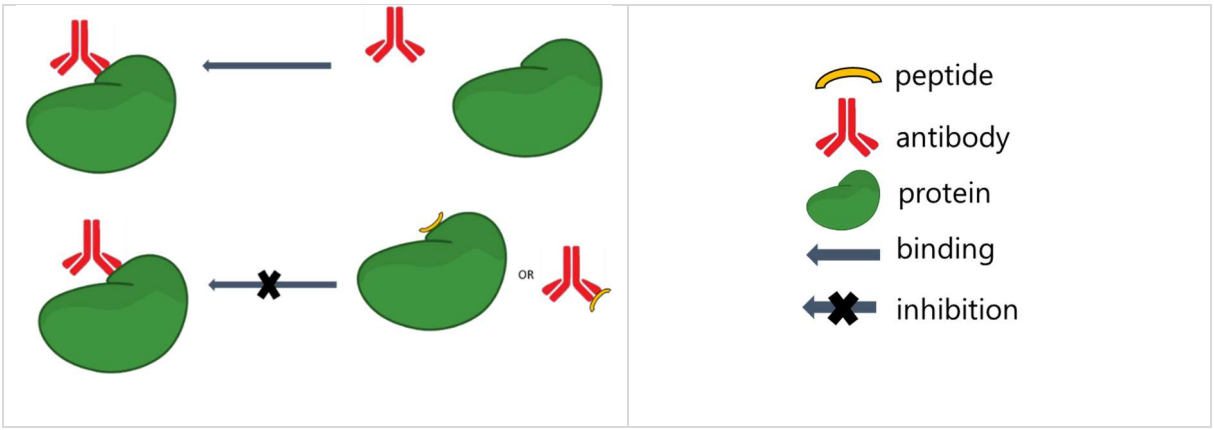
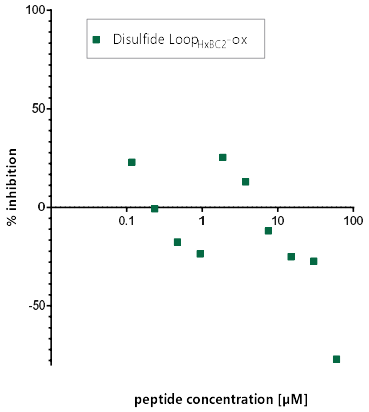


Figure 46: Scheme of the principle of antibody binding to a target protein (antigen) and inhibition of the interaction by peptides either binding to antibody itself or the antibody´s epitope within the target protein.

Depending on the mechanism, inhibitory activity is driven by either mimicking the target epitope or the paratope region of the antibody. It had been shown that peptides can be powerful binding site or epitope mimics of protein-protein interactions resulting in inhibition up to complete suppression of the superordinate biological effect in the context of viral infections [6, 123]. Epitope mimics of antigens are also called mimotopes.

The goal was to identify peptides that inhibit the interaction of HIV-1 neutralizing antibodies with their antigen by detecting peptides that bind to viral proteins (gp41/gp120). For the identification of new specific peptides that inhibit a designated antibody-antigen interaction, screening or structure-based approaches can be used. Briefly, the structure-based approach

usually uses available 3D structures of the protein-protein interaction of interest. Based on the structure, interaction sites can be determined on the amino acid level. These interaction sites can be displayed in peptides, which are then tested for their ability to inhibit the selected interaction [6]. The goal was to identify peptides that inhibit the interaction of HIV-1 neutralizing antibodies with their antigen by detecting peptides that bind cellular receptors or viral proteins (gp41/ gp120). A promising candidate was the epitope peptide Disulfide $Loop_{HXBc2}$-ox, representing the epitope of the HIV-1 antibody F240 [124]. The epitope mimicking peptide did not show the expected inhibition but rather enhanced the HIV-1 infection (Figure 47). An unexpected enhancement of HIV-1 infection is also known for a peptide, which mimics the discontinuous CD4-binding site [125].



Disulfide $Loop_{HxBc2}$-ox	Bio-XKDQQLLGIWG[CSGKLIC]TTAV-NH_2

[ ] intramolecular disulfide bridge

Figure 47: HIV-1_{NL4-3} inhibition by Disulfide $Loop_{HXBc2}$-ox peptide, tested in SEAP assay.

Since rational structure-based approaches for both targets, CD4 binding site of gp120 and the disulfide loop region gp41, have failed, a screening approach using combinatorial libraries was pursued. Combinatorial libraries are synthetically generated libraries that allow the screening of compounds with a defined set of mixtures, which can be applied to a variety of bioassays. Libraries of peptides up to the decapeptide length have been generated, as well as libraries with small molecule scaffolds (cf. 1.4) [95].

The aim was to screen for peptides that interact with the CD4-binding site of gp120 and the disulfide loop of gp41 respectively. An established concept was optimized to identify inhibitory peptides against HIV-1 antibodies with these two particular epitopes (CD4 binding site and

disulfide loop gp41). The screening was performed in competitive antibody-protein assay systems. The HIV-1 antibodies of interest were the CD4 binding site antibodies: VRC01, VRC03, b12 and F240, which is targeting the disulfide loop region of gp41 (c.f. 1.2.3). The screening approach used mixture-based peptide libraries. It offers the advantageous opportunity to simultaneously test millions of compounds that are arranged in organized mixtures (cf. 1.4). At first, scaffold ranking samples that include all compounds in one sample were screened to identify active peptide library samples. This was followed by positional scanning of a selected library and completed by testing the deconvoluted individual peptides in $gp120_{HXBc2}$ -VRC01 competitive assay. The performed assay optimization prior to the screening included the determination of signal to noise ratio, verification of reproducibility and introduction of positive controls to guarantee robust assay conditions with a low variability. The competitive assays were based on the binding of b12, VRC01 and VRC03 to $gp120_{HXBc2}$ and the binding of F240 to $gp41_{HXBc2}$. All tested library samples were kindly provided by TPIMS (Florida) and the scaffold ranking as well as the majority of the positional scanning testing were performed during a research stay in the laboratory of Dr. Pinilla at TPIMS (California).

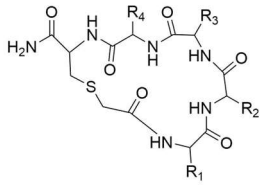
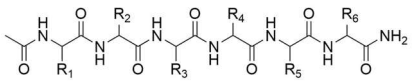
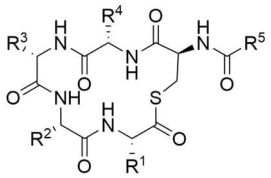
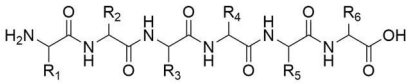
4.2.1 Scaffold Ranking Testing to Identify Active Peptide Libraries

In order to identify the most promising library for each antibody-protein interaction, scaffold ranking samples of mixture-based combinatorial peptide libraries were tested with established competitive antibody assays. The scaffold ranking sample (*All X sample*) presented the entire library in one sample. The available scaffold ranking set, consisting of scaffold ranking samples of 18 different libraries, were tested in four to five concentrations depending on the antibody-protein competitive assay. The libraries are named TPI followed by a four-digit code to identify the library.

For the F240-$gp41_{HXBc2}$ assay five concentrations of scaffold ranking samples were tested. Due to the previously determined signal to noise ratio of the F240-$gp41_{HXBc2}$ assay, only inhibition higher than 35 % was considered for the activity evaluation of the scaffold ranking samples. Inhibition higher than 50 % was achieved by four samples in higher concentrations (2.00 mg/mL, 1.25 mg/mL and 0.63 mg/mL). The other samples only showed moderate to low

inhibitory activity. Therefore, the four scaffold samples of TPI 2068, 2119, 2320 and 2321 had an outstanding higher inhibitory activity compared to the rest of the tested samples (Table 28). These four samples can be sorted into two groups: cyclic peptide (TPI 2320 and 2321) and L-AA hexapeptide (TPI 2068 and 2119) library samples.

Table 28: Scaffolds of the most active libraries which inhibited the F240-gp41$_{HXBc2}$interaction.

TPI 2320: Cyclic thio peptide 	TPI 2068: Hexapeptide (L-AA) acetylated, amide 
TPI 2321: Cyclic peptide 	TPI 2119: Hexapeptide (L-AA) free amine, carboxyl 

The best concentration dependent activity was observed for TPI 2321 and TPI 2068. The two active cyclic peptide libraries showed inhibition higher than 60 % at 1.25 mg/mL. Out of the L-AA hexapeptide library group inhibition higher than 50 % was achieved by the scaffold samples of TPI 2219 at 0.63 mg/mL and TPI 2068 at 2.0 mg/mL. The other tested L-AA hexapeptide libraries were almost inactive, as well as the tested tetra- and decapeptide library samples (Table 29). Based on these results, the most promising candidates for a follow up by positional scanning were TPI 2119 and TPI 2321.

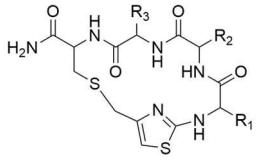
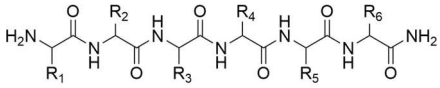
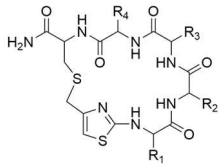
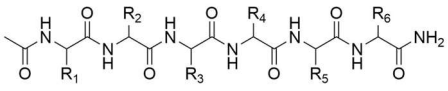
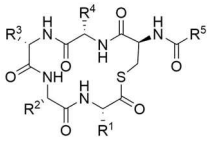
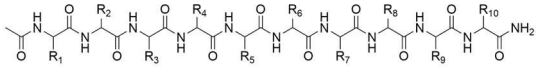
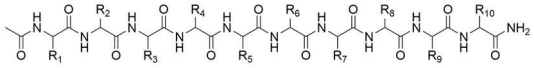
Table 29: Inhibition of the F240-gp41$_{HXBc2}$ interaction by the scaffold ranking sample set, color code for each concentration, lowest value = red, quantile = yellow, highest value = green, tested in Competitive Assay Format A.

	F240-gp41$_{HxBC2}$ assay	concentration [mg/mL]				
		2.00	**1.25**	**0.63**	**0.31**	**0.16**
TPI	structure	AVG [% inhibition]				
1988	Cyclictripeptide thiazole	45	36	15	16	9
1989	Cyclictetrapeptide thiazole	33	30	20	9	10
2040	Decapeptide (L-AA) acetylated, amide	n.d.	31	27	26	21
2068	Hexapeptide (L-AA) acetylated, amide	56	45	36	20	25
2069	Hexapeptide (L-AA) free amine, amide	11	17	18	21	25
2085	Hexapeptide (L-AA) acetylated, carboxyl	26	27	16	26	8
2119	Hexapeptide (L-AA) free amine, carboxyl	54	56	53	47	26
2225	N-naphthyl-dipeptide	17	23	14	22	27
2226	Dipeptide	13	29	8	17	26
2228	Acylated dipeptide	39	43	30	24	26
2320	Cyclic thio tetrapeptide	67	64	50	19	17
2321	Cyclic peptide	60	66	60	35	33
2337	Hexapeptide (D-AA) free amine, amide	12	4	7	15	0
2338	Hexapeptide (D-AA) acetylated, amide	41	29	28	27	9
2339	Tetrapeptide (D-AA) free amine, amide	10	15	10	18	9
2340	Tetrapeptide (D-AA) acetylated, amide	23	17	15	16	6
2500	Tripeptide (L-, D-AA) free amine, amide	8	5	7	14	1
2512	Decapeptide w/citrulline (L-AA) acetylated, amide	34	26	22	34	12

n.d. = not determined

A slightly modified scaffold ranking set was screened in gp120$_{HXBc2}$-b12 competitive assay. In contrast to the results obtained with F240-gp41$_{HXBc2}$ interaction, inhibition up to 89 % was achieved. The level of differentiation between the tested *all X samples* was low at 2.00 mg/mL and 1.25 mg/mL, because half of the samples showed inhibition values higher than 50 %. For this reason, the activity analysis was focused on the three lower concentrations. The inhibitory activity allowed the sorting into high, medium and low active samples. The high active library samples could be divided in two groups: cyclic and linear scaffolds (Table 30).

Table 30: Scaffolds of the most active libraries which inhibited the $gp120_{HXBc2}$-b12 interaction.

TPI 1988: Cyclictripeptide thiazole 	TPI 2337: Hexapeptide (D-AA) free amine, amide 
TPI 1989: Cyclictetrapeptide thiazole 	TPI 2338: Hexapeptide (D-AA) acetylated, amide 
TPI 2321: Cyclic peptide 	TPI 2040: Decapeptide (L-AA) acetylated, amide 
	TPI 2512: Decapeptide (L-AA) free amine, amide 

One group consisted of cyclic peptide scaffolds (TPI 1988, 1989, 2321). The cyclic thiazole scaffolds (TPI 1988 and TPI 1989) and the thioether scaffold (TPI 2321) showed inhibition around 60 – 70 % at 0.63 mg/mL and a good concentration response. The second group of high active samples was more heterogenous. The D-AA hexapeptide library (TPI 2338) was the most potent inhibitor of this group with 68 % inhibition at 0.625 mg/mL. The decapeptide (TPI 2512) and the other D-AA hexapeptide samples (TPI 2337 and TPI 2040) were able to inhibit around 40 to 50 % at 0.625 mg/mL with an overall good concentration dependency of all samples (Table 31). Based on these results, TPI 2321 and TPI 2338 were the most promising candidates for identifying inhibitory compounds after positional scanning and deconvoluting individual peptides.

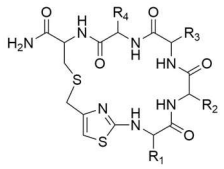
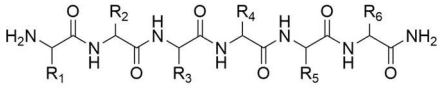
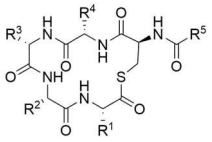
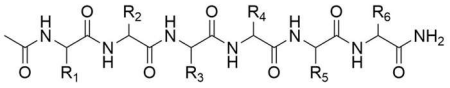
Table 31: Inhibition of the $g120_{HXBc2}$-b12 interaction by the scaffold ranking sample set, color code for each concentration, lowest value = red, quantile = yellow, highest value = green, tested in Competitive Assay Format B (at least in duplicates).

	b12-$gp120_{HxBC2}$ assay	concentration [mg/mL]				
		2.00	1.25	0.63	0.31	0.16
TPI	structure	AVG [% inhibition]				
1988	Cyclictripeptide thiazole	89	77	72	55	39
1989	Cyclictetrapeptide thiazole	82	80	70	49	24
2040	Decapeptide (L-AA) acetylated, amide	n.d.	58	44	22	15
2068	Hexapeptide (L-AA) acetylated, amide	71	63	43	35	0
2069	Hexapeptide (L-AA) free amine, amide	50	35	30	32	13
2085	Hexapeptide (L-AA) acetylated, carboxyl	36	36	-16	13	23
2119	Hexapeptide (L-AA) free amine, carboxyl	54	50	32	18	18
2225	N-naphthyl-dipeptide	32	22	6	20	-8
2226	Dipeptide	27	34	10	13	7
2228	Acylated dipeptide	43	41	27	22	16
2320	Cyclic thio tetrapeptide	87	78	58	11	14
2321	Cyclic peptide	63	91	83	56	39
2337	Hexapeptide (D-AA) free amine, amide	73	69	49	42	25
2338	Hexapeptide (D-AA) acetylated, amide	87	81	68	55	39
2339	Tetrapeptide (D-AA) free amine, amide	58	43	24	13	6
2340	Tetrapeptide (D-AA) acetylated, amide	54	37	17	9	10
2500	Tripeptide (L-, D-AA) free amine, amide	53	23	7	n.d	n. d
2512	Decapeptide w/citrulline (L-AA) acetylated, amide	65	58	49	35	26

n.d. = not determined

Furthermore, CD4 binding site antibody assays were screened with the scaffold ranking set. These two antibodies, called VRC01 and VRC03, belong to the same family of broadly HIV-1 neutralizing antibodies. For each one of these antibodies an individually optimized competitive assay system was established. Up to three different concentrations per scaffold sample were tested in $gp120_{HXBc2}$-VRC01 competitive assay. At 1.25 mg/mL six scaffold samples achieved inhibition higher than 60 %. Taking a closer look at the concentration dependency of these six samples, five out of six were able to maintain inhibition around 50 % at 0.63 mg/mL. These high active samples can be sorted into two scaffold groups: cyclic peptides and hexapeptides (Table 32).

Table 32: Scaffolds of the most active libraries which inhibited the $gp120_{HXBc2}$-VRC01 interaction.

TPI 1989: Cyclictetrapeptide thiazole 	TPI 2337: Hexapeptide (D-AA) free amine, amide 
TPI 2321: Cyclic peptide 	TPI 2338: Hexapeptide (D-AA) acetylated, amide 

The cyclic scaffolds (TPI 1989 and TPI 2321) had a good concentration dependency with a small lead in activity compared to the other groups of active scaffold samples at 0.63 mg/mL. TPI 2337 and TPI 2338, both having a D-AA hexapeptide scaffold, were the second most active samples and were followed by TPI 2068, a L-AA hexapeptide sample (Table 33). In particular, the D-AA hexapeptide scaffold samples were more active than most of the L-AA hexapeptide scaffold samples, as well as all other tested linear scaffold samples. Neither decapeptide nor tetrapeptide samples showed high effective inhibition of VRC01-$gp120_{HXBc2}$ interaction. Both D-AA hexapeptide libraries (TPI 2337 and TPI 2338), TPI 1989 and TPI 2321 are pursuable libraries for further screening and identification of active individual compounds.

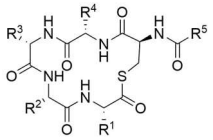
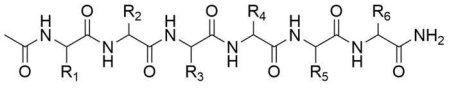
Table 33: Inhibition of the VRC01-g120$_{HXBc2}$ interaction by the scaffold ranking sample set, color code for each concentration, lowest value = red, quantile = yellow, highest value = green, tested in Competitive Assay Format B (at least in duplicates).

	VRC01-gp120$_{HxBC2}$ assay	concentration [mg/mL]			
		1.25	0.75	0.63	0.31
TPI	structure	AVG [% inhibition]			
1988	Cyclictripeptide thiazole	68	n.d.	n.d.	n.d.
1989	Cyclictetrapeptide thiazole	70	n.d.	60	36
2040	Decapeptide (L-AA) acetylated, amide	60	50	38	n.d.
2068	Hexapeptide (L-AA) acetylated, amide	51	n.d.	49	36
2069	Hexapeptide (L-AA) free amine, amide	17	n.d.	23	34
2085	Hexapeptide (L-AA) acetylated, carboxyl	-32	n.d.	45	n.d.
2119	Hexapeptide (L-AA) free amine, carboxyl	59	43	33	n.d.
2225	N-naphthyl-dipeptide	7	n.d.	11	10
2226	Dipeptide	15	n.d.	27	26
2228	Acylated dipeptide	25	n.d.	26	31
2320	Cyclic thio tetrapeptide	49	n.d.	6	28
2321	Cyclic peptide	78	65	66	n.d.
2337	Hexapeptide (D-AA) free amine, amide	57	52	52	n.d.
2338	Hexapeptide (D-AA) acetylated, amide	63	61	52	n.d.
2339	Tetrapeptide (D-AA) free amine, amide	31	29	n.d.	n.d.
2340	Tetrapeptide (D-AA) acetylated, amide	36	34	n.d.	n.d.
2512	Decapeptide w/citrulline (L-AA) acetylated, amide	40	32	16	n.d.

n.d. = not determined

The second antibody of the VRC01 family is called VRC03 and was also screened with the scaffold ranking set. For this reason, a competitive ELISA based on gp120$_{HXBc2}$-VRC03 interaction was established and used for library sample testing. The activity profile and level of differentiation within the tested scaffold samples was quite different compared to the results obtained in VRC01-gp120$_{HXBc2}$ competitive assay. Two samples, cyclic thioether scaffold (TPI 2321) and D-AA hexapeptide (TPI 2338), showed outstanding inhibition up to 80 % with a good concentration response compared to all other tested samples in gp120$_{HXBc2}$-VRC03 competitive assay.

Table 34: Scaffolds of the most active libraries which inhibited the VRC03-gp120$_{HXBc2}$interaction.

TPI 2321: Cyclic peptide	TPI 2338: Hexapeptide (D-AA) acetylated, amide
	

This was followed by another moderately active cyclic scaffold (TPI 2320) with 60 % inhibition at 0.625 mg/mL (Table 35). Both outstanding active libraries (TPI 2320 and TPI 2338) were selected as candidates for further screening of a designated library. Regardless of the different activity profile obtained with VRC01 and VRC03 assay, the two most active samples for VRC03 are also among the four most active library samples for VRC01.

Table 35: Inhibition of the VRC03-g120$_{HXBc2}$ interaction by scaffold ranking samples, color code for each concentration, lowest value = red, quantile = yellow, highest value = green, tested in Competitive Assay Format B (at least in duplicates).

	VRC03-gp120$_{HxBC2}$ assay	concentration [mg/mL]			
		1.25	0.75	0.63	0.31
TPI	structure	AVG [% inhibition]			
1988	Cyclictripeptide thiazole	13	n.d.	n.d.	n.d.
1989	Cyclictetrapeptide thiazole	53	n.d.	44	22
2040	Decapeptide (L-AA) acetylated, amide	52	48	42	n.d.
2068	Hexapeptide (L-AA) acetylated, amide	56	n.d.	50	35
2069	Hexapeptide (L-AA) free amine, amide	17	n.d.	24	31
2085	Hexapeptide (L-AA) acetylated, carboxyl	23	n.d.	52	n.d.
2119	Hexapeptide (L-AA) free amine, carboxyl	55	43	52	n.d.
2225	N-naphthyl-dipeptide	-2	n.d.	18	12
2226	Dipeptide	1	n.d.	31	16
2228	Acylated dipeptide	25	n.d.	30	28
2320	Cyclic thio tetrapeptide	62	n.d.	60	56
2321	Cyclic peptide	80	72	75	n.d.
2337	Hexapeptide (D-AA) free amine, amide	40	36	38	n.d.
2338	Hexapeptide (D-AA) acetylated, amide	71	65	62	n.d.
2339	Tetrapeptide (D-AA) free amine, amide	39	30	n.d.	n.d.
2340	Tetrapeptide (D-AA) acetylated, amide	47	41	n.d.	n.d.
2500	Tripeptide (L-, D-AA) free amine, amide	n.d.	n.d.	24	n.d.
2512	Decapeptide w/citrulline (L-AA) acetylated, amide	23	26	28	n.d.

n.d. = not determined

Based on the availability of the libraries the follow up candidate was selected out of the linear peptide scaffold group. The previous evaluation for each antibody individually revealed that the hexapeptides were more potent than shorter linear peptide scaffold samples. In addition, the decapeptide scaffolds were mostly similar or less active as the hexapeptides and thus main chain prolongation up to ten amino acids was not beneficial for inhibitory activity of the library scaffold sample.

Due to the consistent activity of hexapeptide scaffolds against all three CD4 binding site antibodies, these libraries were considered to be promising candidates for follow up testing with possible deconvolution of individual peptides. Therefore, hexapeptide scaffold results of b12, VRC01 and VRC03 were compared at two concentrations to select follow up candidates. The tested D-AA scaffolds differ in their N- and C-terminal modification. Out of the group of D-AA scaffolds TPI 2338 (acetylated N-terminus) was more potent than TPI 2337 (non-acetylated N-terminus), although the difference was not very distinct for VRC01. Furthermore TPI 2338 showed higher inhibition than all tested L-AA hexapeptide scaffold samples. The L-AA scaffold, TPI 2068, had the best breath of inhibition of all tested L-AA hexapeptide scaffold samples against all three antibodies whereas TPI 2069 and TPI 2085 showed the weakest activity. Although TPI 2119 with acetylated N-terminus and C-terminal acid was the second most potent L-AA scaffold sample. The comparative analysis showed that the D-AA scaffold sample of TPI 2338 was the most promising candidate out of the group of hexapeptides scaffolds, which showed a concentration dependent inhibition in gp120 targeting antibody assays (Table 36).

Table 36: Inhibition of the CD4-binding-site-antibody-g120$_{HXBc2}$ interactions by hexapeptide scaffold ranking samples, tested in Competitive Assay Format B. color code for % inhibition: green: > 50, yellow = 30-50 and red: < 30.

% inhibition
≤ 30
30-50
≥ 50

		VRC01		VRC03		b12	
		library sample concentration [µg/mL]					
		1.25	0.625	1.25	0.625	1.25	0.625
TPI	strucure	AVG [% inhibition]					
2068	Hexapeptide (L-AA) acetylated, amide	51	49	56	50	63	43
2069	Hexapeptide (L-AA) free amine, amide	17	23	17	24	35	30
2085	Hexapeptide (L-AA) acetylated, carboxyl	-32	45	23	52	36	-16
2119	Hexapeptide (L-AA) free amine, carboxyl	59	33	55	52	50	32
2337	Hexapeptide (D-AA) free amine, amide	57	52	40	38	69	49
2338	Hexapeptide (D-AA) acetylated, amide	63	52	71	62	81	68

The scaffold ranking approach can be considered to be successful for all four used antibody assays because it was possible to identify active samples. The performed screening proved the suitability of the scaffold ranking approach to identify active peptide libraries in HIV-1 antibody-protein competitive assays. Moreover, this method was applied to assay systems using HIV-1 neutralizing antibodies targeting the CD4 binding site of gp120 and the disulfide loop of gp41 for the first time. The concentration dependent testing of 18 peptide scaffold samples was the basis for a selection of a promising candidate for further investigation. Having the best concentration activity profile in all three CD4 antibody assays, the hexapeptide library TPI 2338 (D-AA acetylated, amide) was consequently chosen for further investigation up to the deconvolution of individual peptides.

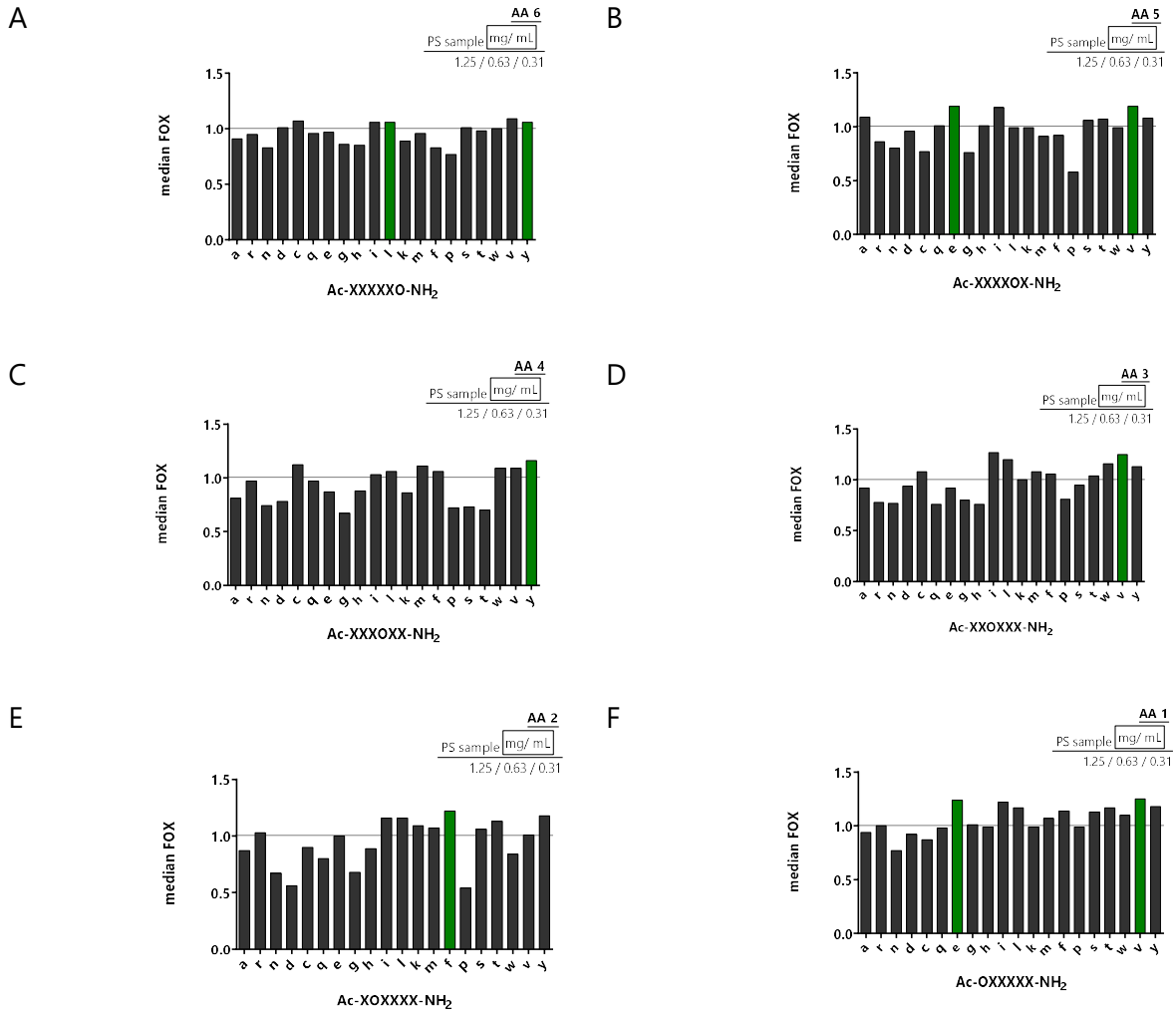
4.2.2 Deconvoluting Individual D-AA Hexapeptides from a PS-SCL which Inhibit gp120$_{HXBc2}$-VRC01 Interaction

The initially performed scaffold ranking led to the identification of active library samples. The subsequent step was the positional scanning of a designated library. Despite the fact that all tested antibody assays revealed active scaffold ranking samples, only the follow up on gp120$_{HXBc2}$-VRC01 competitive assay is presented. The antibody-protein interaction with the

strongest HIV-1 neutralizing potency out of the tested CD4 binding site antibody group was selected for positional scanning. HIV-1 neutralizing antibodies can be sorted by their breadth of inhibiting HIV-1 strains. From the set of tested HIV-1 antibodies VRC01 has the broadest capacity to inhibit HIV-1 infection [29]. Nevertheless, promising activity was detected for all initially screened antibody-protein interactions. The follow up with TPI 2338 was based on its promising activity, as well as the availability of the PS (positional scanning) sample sets (c.f. 4.2.1). TPI 2338 is a positional scanning synthetic library (PS-SCL), which consisted of a D-AA hexapeptide scaffold (C-terminal amide and N-terminally acetylated) and was generated by incorporating twenty different amino acids at each chain position, which results in 64 million individual peptides in total. Each mixture of the positional scanning sample library has one defined amino acid position and multiple variable positions are composed of a mixture of amino acids. Subsequently 20 positional scanning mixtures per position were tested. The purpose of the positional scanning was to use the obtained profiles to deconvolute individual active peptides. The testing of the corresponding positional scanning set of the library was performed at three concentrations with $gp120_{HXBc2}$-VRC01 competitive assay in duplicates. The positional scanning set was composed of 120 mixtures with one defined position. For each concentration and replicate the corresponding scaffold ranking sample (*all X sample*) that initially was used for screening, was included (c.f. 4.2.1). The analysis of the results was not based on percentage of inhibition but instead fold over X (FOX) was applied. In this case the *all X sample* (scaffold sample) was defined as the baseline (X). The decision to choose fold over X was based on two observations: the lack of significant differentiation in activity between the mixtures and the big variation of *all X samples* activity between the replicates. Based on FOX the inhibitory activity of the positional scanning mixtures is ranked by higher or lower inhibition compared to the *all X sample*. FOX was calculated for each replicate and the median FOX equally represented all performed replicates at three different concentrations. The positional scanning profile (20 samples per position) of each position is evaluated separately.

The obtained positional scanning profile for position 6 was not highly differentiated. Only five mixtures showed a median FOX higher than 1 and only in the FOX range of 1.06-1.09. This increase is notably small and it can be concluded that position 6 is not remarkably contributing to the observed activity of the *all X sample*. The samples with a median FOX higher than 1 were further evaluated. One aliphatic representative out of the group of active mixtures with defined

leucine, valine and isoleucine was chosen for deconvolution. Cysteine was excluded due to its unknown redox state in the tested mixtures. Two AAs were chosen for deconvolution out of the group with a median FOX >1 (tyrosine and leucine) (Figure 48 A). The positional scanning profile of position 5 was more differentiated than the one of position 6. A median FOX higher than 1,18 was achieved by three mixtures defined with two aliphatic AAs (isoleucine and valine) and glutamic acid (Figure 48 B). Valine and glutamic acid were selected for deconvolution of position 5. The evaluation of the profile of position 4 revealed that one mixture had a higher median FOX compared to all others. For position 4 only the mixture with defined tyrosine showed a median FOX higher than 1.15 and was consequently selected (Figure 48 C). Moreover, the majority of the samples had a median FOX below 1. Compared to the other positions the median FOX based scanning profile of position 3 had a higher level of differentiation. Mixtures with defined aliphatic AAs (valine, leucine and isoleucine) represented the most active PS samples for position 3 (median FOX >1.2). The median FOX of all three samples was slightly higher compared to all other tested mixtures belonging to this position. Thus, only these three were considered for deconvolution. In this case valine was selected as a representative of the three aliphatic AAs (Figure 48 D). A noteworthy aspect of the scanning profile of position 2 was the high proportion of samples with median FOX values below 0.6. At the upper end of the profile the mixture defined with phenylalanine had a slightly higher median FOX than the following samples and was subsequently chosen for deconvolution (Figure 48 E). The results of the positional scanning profile of position 1 showed the highest activity compared to the other position profiles. The majority had a median FOX equal or higher than 1. The most potent mixtures had either aliphatic AAs (valine and isoleucine) or glutamic acid defined. Therefore, valine and glutamic acid were selected (Figure 48 F). The use of median FOX approach enabled the use of these specific positional scanning profiles for the deconvolution of 24 individual peptides (Figure 48 G).



G

Deconvolution

	N-Terminus					C-Terminus	
Position	1	2	3	4	5	6	
	v	f	v	y	v	l	
	e			m	e	y	
				w			
	2	1	1	3	2	2	**24 peptides**

Figure 48: Inhibition of the g120$_{HXBc2}$-b12 interaction by mixtures of PS-SCL TPI 2338 (D-AA hexapeptide library) shown as median of fold over x (FOX), tested in Competitive Assay Format B (1.25/0.63/0.31 mg/mL) A: mixtures of position 6 B: mixtures of position 5 C: mixtures of position 4 D: mixtures of position 3 E: mixtures of position 2 F: AA mixtures of position 1. A-F: green labeled bars = chosen for deconvolution. O= defines each subsequent position X = mixture of all AA. G: Summary table of selected amino acids per position for deconvolution.

4.2.3 Individual Peptides from Positional Scanning Profile of a D-AA Hexapeptide Library

Based on the inhibition obtained for the positional scanning profiles, 24 peptides deconvoluted from TPI 2338 were successfully synthesized (c.f. 7.1). These peptides represented all of the possible combinations of the selected amino acids per position. Initially the peptides were tested without any further purification by HPLC. In the following these crude peptides are referred to as VRC01-crude-1 to -24. A careful evaluation of product quality was necessary and signals corresponding to the expected peptide sequences were identified by LC-MS (cf. 7.1). The crude peptides were tested at 200 µM in $gp120_{HXBc2}$-VRC01 competitive assay in duplicates. The relatively high concentration of 200 µM was chosen based on the positional scanning results with a moderate level of inhibition. At the tested concentration only two crude peptides (VRC01-crude-22 and -crude-23) achieved an inhibition higher than 50 %, followed by VRC01-crude-21 with a higher variable but similar activity. The three peptides had a high level of sequence homology with only two positions varying (Figure 49).

VRC01-crude-21	VRC01-crude-22	VRC01-crude-23
Ac-**efvw**vl-NH_2	Ac-**efvw**vy-NH_2	Ac-**efvw**el-NH_2

Figure 49: Sequences of the most active VRC01 crude peptides.

It is likely that this shared sequential motif was contributing to their activity. That is why all three peptides with a set of controls were selected for purification. The set of controls reflected the structural activity relationship (SAR), meaning that only one position differed compared to the chosen lead sequence. The set will be called SAR controls in the following (Figure 50 and Table 37). The aim was to test the influence of only exchanging single amino acids of the active peptides.

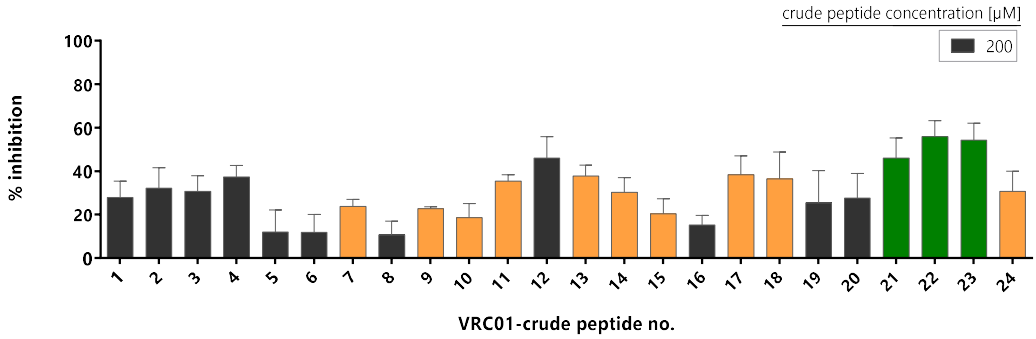


Figure 50: Inhibition of the VRC01-g120$_{HXBc2}$ interaction, tested in Competitive Assay Format D. VRC01-gp120HXBc2 Assay single concentration of VRC01-crude peptides (200µM in assay main buffer), n=6. highlighted bars: chosen for purification: green= high activity yellow = SAR controls.

Each active peptide with the corresponding SAR controls will be presented individually in the following. The purified variant of VRC01-crude-22, called VRC01-ReH12, as well as the corresponding SAR controls were tested in a concentration dependent manner. VRC01-ReH12 and VRC01-ReH11 were the only two hexapeptides with concentration dependent activity up to 70 % inhibition. These two only differed at position 6 (tyrosine to leucine exchange), implying that the contribution of tyrosine is low to the overall activity because replacing the hydroxyphenyl side chain with isobutyl was tolerated. The replacement of valine to glutamic acid at position 5 (VRC01-ReH14), tryptophan to either methionine or tyrosine at position 4 (VRC01-ReH10 and -ReH7) and glutamic acid at position 1 (VRC01-ReH4) resulted in a loss of inhibition (Figure 51 A and B).

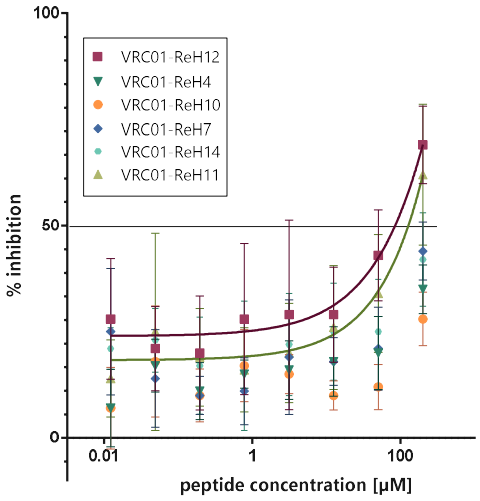
Table 37: Selected VRC01-crude peptide sequences and their corresponding SAR controls, highlighted in green = lead sequence, bold = position that is different compared to the lead peptide.

VRC01-crude-21	Ac-efvwvl -NH_2
VRC01-crude-9	Ac-**v**fvwvl -NH_2
VRC01-crude-17	Ac-efv**m**vl -NH_2
VRC01-crude-13	Ac-efv**y**vl -NH_2
VRC01-crude-23	Ac-efvwe**l** -NH_2
VRC01-crude-22	Ac-efvwv**y** -NH_2

VRC01-crude-22	Ac-efvwvy -NH_2
VRC01-crude-10	Ac-**v**fvwvy -NH_2
VRC01-crude-18	Ac-efv**m**vy -NH_2
VRC01-crude-14	Ac-efv**y**vy -NH_2
VRC01-crude-24	Ac-efvwe**y** -NH_2
VRC01-crude-21	Ac-efvwv**l** -NH_2

VRC01-crude-23	Ac-efvwel -NH_2
VRC01-crude-11	Ac-**v**fvwel -NH_2
VRC01-crude-7	Ac-vfv**m**el -NH_2
VRC01-crude-15	Ac-efv**y**el -NH_2
VRC01-crude-21	Ac-efvw**v**l -NH_2
VRC01-crude-24	Ac-efvwe**y** -NH_2

A



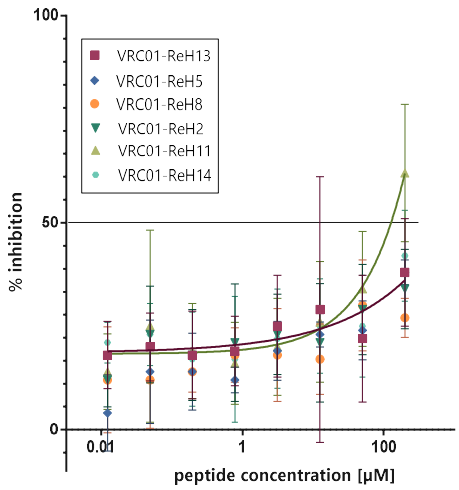
B

	sequence							
name		1	2	3	4	5	6	
VRC01-ReH12	Ac-	e	f	v	w	v	y	NH_2
VRC01-ReH4	Ac-	v	f	v	w	v	y	NH_2
VRC01-ReH10	Ac-	e	f	v	m	v	y	NH_2
VRC01-ReH7	Ac-	e	f	v	y	v	y	NH_2
VRC01-ReH14	Ac-	e	f	v	w	e	y	NH_2
VRC01-ReH11	Ac-	e	f	v	w	v	l	NH_2

Figure 51: Inhibition of the VRC01-g120$_{HXBc2}$ interaction by VRC01-ReH12 (purified variant of VRC01-crude-22) and its corresponding SAR controls, tested in Competitive Assay Format D A: Concentration dependent inhibition curves of VRC01-ReH peptides. B: Table with peptide sequences and highlighted single mutations of the SAR controls.

This indicated that maintaining glutamic acid at position 1 was important and that an introduction of an additional negative charge at position 5 was not beneficial for activity. The second chosen active crude peptide was VRC01-crude-23 with the corresponding purified variant VRC01-ReH13. The purified peptide showed a low inhibition curve below 50 %. The inhibition of the crude peptide was not reproduced by the purified one, and therefore a detailed analysis of the controls was not necessary. The only active SAR control of VRC01-ReH13 was VRC01-ReH11, which was the second active lead peptide (Figure 52).

A



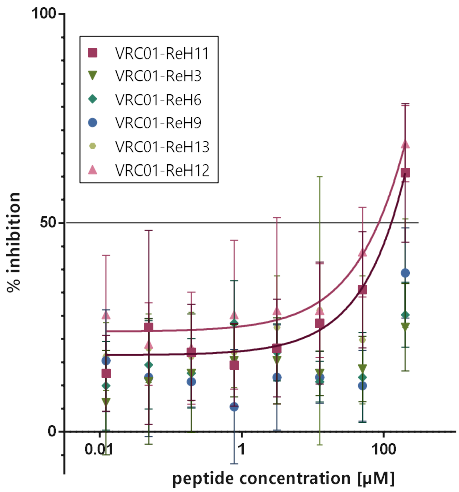
B

	sequence							
		1	2	3	4	5	6	
VRC01-ReH13	Ac-	e	f	v	w	e	l	NH_2
VRC01-ReH5	Ac-	v	f	v	w	e	l	NH_2
VRC01-ReH8	Ac-	e	f	v	y	e	l	NH_2
VRC01-ReH2	Ac-	v	f	v	m	e	l	NH_2
VRC01-ReH11	Ac-	e	f	v	w	v	l	NH_2
VRC01-ReH14	Ac-	e	f	v	w	e	y	NH_2

Figure 52: Inhibition of the VRC01-$g120_{HXBc2}$ interaction by VRC01-ReH13 (purified variant of VRC01-crude-23) and its SAR controls, tested in Competitive Assay Format D. A: Concentration dependent inhibition curves of VRC01-ReH peptides. B: Table with peptide sequences and highlighted single mutations of the SAR controls.

The controls of VRC01-ReH11, the corresponding purified peptide of VRC01-crude-21, confirmed the results obtained with VRC01-ReH12 (Figure 53). The replacement of glutamic acid (position 1, VRC01-ReH5), tryptophan (position 4, VRC01-ReH2 and -ReH8) and valine (position 5, VRC01-ReH11) with an amino acid with different chemical character resulted in severe activity loss up to inactivity. That is why an active 5mer peptide sequence can be proposed: Ac-efvwvO-NH_2, in which O can be occupied by a range of different amino acids.

A



B

		sequence						
		1	2	3	4	5	6	
VRC01-ReH11	Ac-	e	f	v	w	v	l	NH_2
VRC01-ReH3	Ac-	v	f	v	w	v	l	NH_2
VRC01-ReH9	Ac-	e	f	v	m	v	l	NH_2
VRC01-ReH6	Ac-	e	f	v	y	v	l	NH_2
VRC01-ReH13	Ac-	e	f	v	w	e	l	NH_2
VRC01-ReH12	Ac-	e	f	v	w	v	y	NH_2

Figure 53: Inhibition of the VRC01-$g120_{HXBc2}$ interaction by VRC01-ReH11 (purified variant of VRC01-ReH21) and its SAR controls, tested in Competitive Assay Format D. A: Concentration dependent inhibition curves of VRC01-ReH peptides. B: Table with peptide sequences and highlighted single mutations of the SAR controls.

The testing of purified hexapeptides derived from TPI 2338 led to two active peptides (VRC01-ReH11 and -ReH12) which vary only in position 6 of the peptide sequence. The individual peptides reconfirmed the low specificity of position 6 already observed during screening of positional scanning samples. The profile of position 6 had only a small benefit for activity compared to the initially tested all X sample (c.f. 4.2.1). In the tested concentration range from 0.01 – 200 µM neither of the tested peptides reached 100 % inhibition of the $gp120_{HXBc2}$-VRC01 interaction and the general inhibitory activity was too low to use IC_{50} values for activity comparison. Considering this fact, another parameter was calculated for evaluation: the IC_{30}. The applied IC_{30} represents the 30 % inhibitory concentration per peptide and was determined based on the inhibition results of the purified peptides. The IC_{30} allowed the numerical ranking of the inhibitory activity of the peptides in an adequate activity range. Based on the ranking by IC_{30}, VRC01-ReH12 was found to be the most active peptide. The second identified active peptide, VRC01-ReH11, had a six-fold lower IC_{30}, but still showed considerable

activity. The other peptides showed a decreased activity compared to VRC01-ReH11 from nearly no activity to inactivity (Table 38). The decrease in activity due to changing a single amino acid position of the active lead peptide (VRC01-ReH11) was elucidated in detail. The replacement of tyrosine with leucine at position 1 yielded in decrease in activity whereas any other positional change was not tolerated at all. It can be concluded that the activity of the lead peptide is expected to be sequence-dependent.

Table 38: VRC01-ReH peptides sorted by IC_{30} calculated from the obtained inhibition curves in $gp120_{HXBc2}$-VRC01 assay (Competitive Assay Format B), green = single AA substitution variants of VRC01-ReH12, bold = most active lead peptide.

Peptide	sequence	P1	P2	P3	P4	P5	P6	IC_{30} [µM]
VRC01-ReH12	**Ac-efvwvy -NH2**	**e**	**f**	**v**	**w**	**v**	**y**	**2**
VRC01-ReH11	Ac-efvwvl -NH2	e	f	v	w	v	l	12
VRC01-ReH5	Ac-vfvwel -NH2	v	f	v	w	e	l	58
VRC01-ReH4	Ac-vfvwvy -NH2	v	f	v	w	v	y	169
VRC01-ReH8	Ac-efvyel -NH2	e	f	v	y	e	l	194
VRC01-ReH2	Ac-vfvmel -NH2	v	f	v	m	e	l	>200
VRC01-ReH14	Ac-efvwey -NH2	e	f	v	w	e	y	>200
VRC01-ReH3	Ac-vfvwvl -NH2	v	f	v	w	v	l	>200
VRC01-ReH10	Ac-efvmvy -NH2	e	f	v	m	v	y	>200
VRC01-ReH1	Ac-vfvyvl -NH2	v	f	v	y	v	l	>200
VRC01-ReH6	Ac-efvyvl -NH2	e	f	v	y	v	l	>200
VRC01-ReH7	Ac-efvyvy -NH2	e	f	v	y	v	y	>200
VRC01-ReH9	Ac-efvmvl -NH2	e	f	v	m	v	l	>200
VRC01-ReH13	Ac-efvwel -NH2	e	f	v	w	e	l	>200

The peptides VRC01-ReH11, -ReH12 and -ReH13 were further tested in different assay systems to evaluate the selectivity of the activity. The ability to inhibit the $gp120_{HXBc2}$-b12 and F240-$gp41_{HXBc2}$ interaction was investigated. VRC01-ReH11 was able to inhibit the $gp120_{HXBc2}$-b12 but not the F240 $gp41_{HXBc2}$ interaction (Figure 54 A). The same pattern was observed for VRC01-ReH12, however almost 50 % inhibition at 200 µM was also detected in F240-$gp41_{HXBc2}$ competitive assay (Figure 54 B). VRC01-ReH13 was even more capable to inhibit the b12-$gp120_{HXBc2}$-interaction at the highest tested concentration (Figure 54 C). The F240-$gp41_{HXBc2}$ interaction was not accessible for inhibition by VRC01-ReH11 but VRC01-ReH12 had a moderate efficacy. VRC01-ReH11 demonstrated its preference for $gp120_{HXBc2}$ targeting antibody interactions but no selectivity between VRC01 and b12 was achieved. To address the question whether the activity in competitive assays is based on binding of the peptide to the antibody or to the proteins ($gp120_{HXBc2}$, $gp41_{HXBc2}$), biotinylated variants of VRC01-ReH11, -ReH12 and -ReH13 were tested in direct binding assays. In the tested concentrations no binding to b12, VCR01 (two-fold excess compared to competitive assay conditions) and $gp41_{HXBc2}$ (five-fold excess compared to competitive assay conditions) was detectable. Only low binding signals with a signal to noise ratio of up to four was obtained in $gp120_{HXBc2}$ binding assays (chapter 7.2 Figure 70). This weak binding strength could indicate that the activity in competitive assays is due to the ability to bind gp120. The overlapping binding epitope of b12 and VRC01 could explain the non-selective of VRC01-ReH11 and -ReH12.

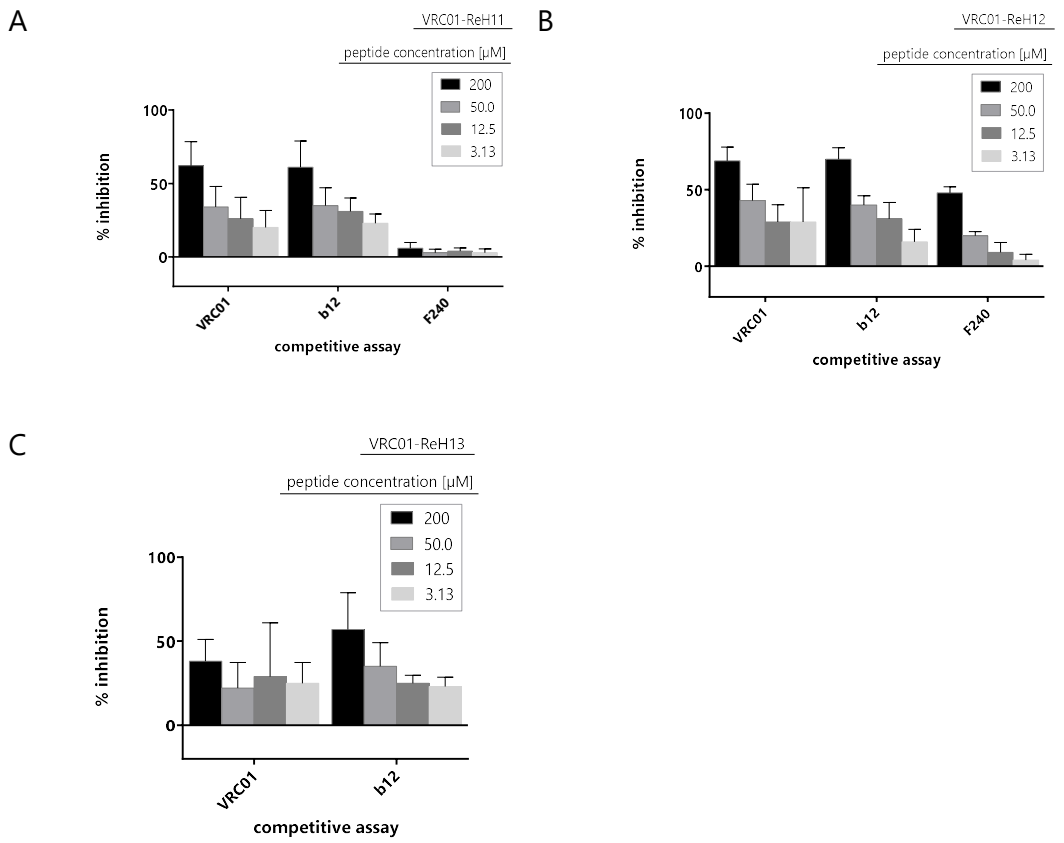


Figure 54: Specificity testing - Inhibition of antibody-protein interaction by VRC011-ReH11, VRC01-ReH12 and VRC01-ReH13 A: Inhibition by VRC01-ReH11 B: Inhibition by VRC01-ReH12 C: Inhibition by VRC01-ReH13 (F240: Competitive Assay Format A, b12/VRC01: Competitive Assay Format D).

L-amino acid variants (VRC01-ReH15, -ReH16 and -ReH17) of VRC01-ReH11, -ReH12 and -ReH13 were synthesized and tested for their ability to inhibit VRC01-gp120$_{HXBc2}$ interaction. The L-AA-variants VRC01-ReH15 (L-AA variant of VRC01-ReH11) and VRC01-ReH17 (L-AA variant of VRC01-ReH13) were inactive. VRC01-ReH16 (L-AA variant of VRC01-ReH12) showed low inhibitory activity below 39 % (Figure 55). Consequently, the inhibitory activity of VRC01-ReH11 and -ReH12 appeared to be selective for the D-AA variant of the peptide sequence.

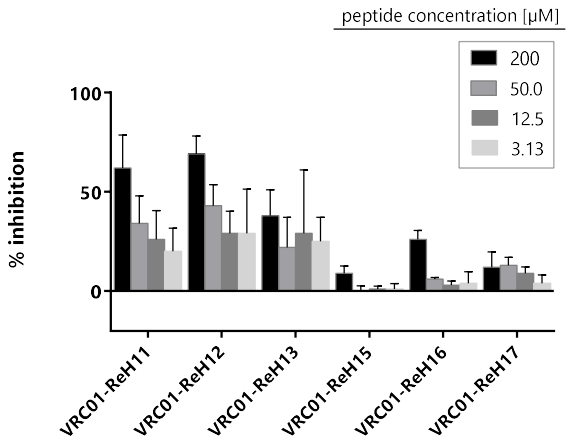


Figure 55: Inhibition of VRC01-gp120$_{HXBc2}$ interaction by VRC01-ReH peptides (including L-AA variants), tested in Competitive Assay Format D.

5 Discussion

5.1 Protein-derived PEGI Peptides Interfering with HIV-1 Entry

5.1.1 PEGI: A Mimic of the E2 Protein Activity?

The HPgV-1 surface proteins E1 and E2, presented as dimers on the particle surface, are postulated to be the determinant of HPgV-1 host cell entry based on recent results and known pathways of related viruses. The precise mechanism remained unclear [65]. Apart from the role of mediating viral entry, several reports exist in literature about their interference with HIV-1 replication (c.f. 1.3.2) [64, 126]. The E2 protein-derived peptides (PEGI and its derivates) that were analyzed and characterized, exhibited effective HIV-1 neutralization in vitro. In a top-down approach starting from the 20mer lead peptide the minimum length of the active sequence was identified. During this process a new shorter and more active peptide PEGI-D35 was developed. The active PEGI peptides represent a sequence from the N-terminal region of the E2 protein, which share a number of unique features: C-terminally located negative charge and three cysteines, preferably in the reduced form. PEGI represents AA 45 to 64 of the E2 protein of the HPgV-1 isolate called Iowa, which belongs to the subtype 2 (c.f. 7.2, Figure 71). Little is known about the secondary structure and disulfide bridge connectivity of the E2 protein. A model proposing that the N-terminus (AA 1-75) remains unstructured had been published [22]. To acquire more information about the surface accessibility, the secondary structure of the E2 protein sequence section AA 1-340, Iowa isolate (Genbank Accession number: AF121950), was predicted by using the tool *predict protein* [127-130]. The simulation showed that parts of the E2 protein represented by the PEGI sequence are buried inside the secondary protein structure as well as partly structured. The buried part includes the C-terminal crucial cysteine, whereas the hydrophobic core motif cannot be specifically defined as buried. Furthermore, the important C-terminal negative charge of the PEGI sequence is, as anticipated, likely to be surface exposed. This prediction supports the assumption that the PEGI sequence is at least partly surface accessible (N-terminus) for binding other entities like proteins. The overall sequence is rendered mostly unstructured and can therefore be assumed to exist in a flexible, linear manner

without stable secondary structure (Figure 56). The results support that the flexible PEGI peptide, representing a section of the N-terminus, could mimic E2 protein-protein interaction.

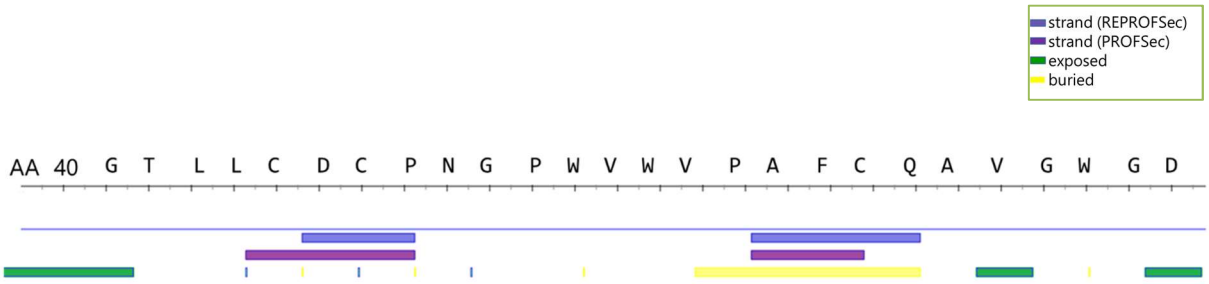


Figure 56: Prediction of secondary structure of section AA40-67 of the E2 protein (AA 1-340, Iowa isolate AF121950) created with *predict protein* [127-130].

The cysteine connectivity is another important PEGI peptide property that needed to be evaluated. The peptide sequence contains three cysteines which are crucial for its activity and especially their redox state. Based on the even number of cysteines (18) in the cytoplasmic domain of the E2 protein it can be speculated that all cysteines are oxidized and part of a disulfide bridge. Despite the fact that proteins with free cysteines are described in the literature, the main role of cysteines in proteins is stabilizing the secondary structures *via* covalent intramolecular connections [105, 131]. To overcome the lack of information about the connectivity of cysteines in a protein sequence of interest, bioinformatic tools are available. The E2 protein (AA 1-340, Iowa isolate) was analyzed *via DiANNA* 1.1 webserver [132-134]. The tool predicted that all cysteines could be connected *via* intramolecular disulfide bridges, but the PEGI peptide cysteine would be connected with cysteines beyond the section represented by the peptide (chapter 7.2, Figure 72). The simulation supports that intramolecular disulfides within the PEGI sequence would rather distort than stabilize the secondary structure of the whole E2 protein. Thus, monomeric PEGI peptides have a higher likelihood to represent a stable E2 protein conformation. The results for the PEGI variants reflected that bridging did not result in an increase of affinity or inhibitory activity. Moreover, the cysteine class prediction tool showed that reduced thiol moieties for the PEGI peptide section scored considerably higher than the oxidized ones (c.f. 7.2, Figure 73). The preference for reduced cysteines in the E2 protein prediction correlated with the results for the PEGI peptides. However, the prediction tools can only provide basic information and detailed structural analysis will be necessary to identify the predominantly formed disulfide bridges of the stable protein. In recent years a paradigm

change occurred in the basic biochemistry of cysteine in proteins. The understanding of the role of cysteines as a factor for protein stability and rigidity was extended to a second functionality: redox-active cysteines. These cysteines can be classified into different types of so-called disulfide switches [135]. The structural analysis of the PEGI sequence exhibited that the functionality of the cysteines is more redox-active than stabilizing secondary structure by forming rigid bridges. An already identified factor for redox-sensitive cysteines in polypeptides is the close proximity to prolines, which is the case for two cysteines in the PEGI sequence [135]. A promising starting point was provided with the peptide analysis. Further characterization of the E2 protein would be required to transfer the hypothesis of the redox-sensitive cysteines from the peptide to the protein level. It can be speculated that the redox activity of the cysteines of PEGI is involved in the mechanism of HIV-1 inhibition by PEGI. The results showed that PEGI variants with reduced cysteines were distinctly more active than oxidized ones and thus supporting this hypothesis.

Another aspect that needed to be assessed for PEGIs possible role as a mimic of E2 activity is the variability of the sequence within the different HPgV-1 isolates. The performed database analysis of available E2 protein sequences revealed that the PEGI sequence is highly conserved and contained only three variable positions at a sequence length of 20 AA. The testing of the strain variants in HIV-1 neutralization assay remarkably demonstrated that the most abundant variants were also the most potent inhibitors in vitro (PEGI, PEGI-D54). The investigated HIV-1 inhibitory activity of PEGI, a protein-derived peptide, is not an isolate specific sequence but a conserved motif among worldwide circulating HPgV-1 strains and its impact on viral interference should not be underestimated. Nevertheless, the E2 protein-derived peptide PEGI is not the only determinant for inhibition of HIV-1 infection by HPgV-1. Beside PEGI, different parts of E2 and E1 as well as peptides derived from non-structural viral proteins of HPgV-1 had been described as active in HIV-1 infection assays [64, 78, 81, 86, 136]. Thus, other pathways causing viral interference *via* up or down regulating of chemokines were proposed [77, 87]. The viral interference between HIV-1 and HPgV-1 is a complex phenomenon with multiple pathways and yet not fully understood. To classify the studied activity of PEGI as a mimic of the E2 protein activity would not correctly reflect the gained understanding. The results were rather providing a possible mechanism for viral interference on the peptide level. The detailed study of the structure activity relationship of PEGI contributed unprecedented insight in the complex

mechanism characterizing the interaction involving the envelope proteins of the human viruses HPgV-1 and HIV-1.

5.1.2 Exploring the Mechanism of HIV-1 Inhibition by PEGI

The ability of PEGI to inhibit HIV-1 infection in vitro is very likely based on a mechanism involving binding to viral ENV. The results of this work indicated that the binding to gp120 is an important determinant for PEGI activity in vitro. ENV undergoes a number of conformational changes during the HIV-1 entry process. Binding to monomeric recombinant gp120 as well as previously demonstrated entry inhibition in virion-based fusion assays are proposing that HIV-1 replication is inhibited by blocking the viral entry process [85]. Additionally, it was shown that the CD4 binding of the virus was not affected in the presence of PEGI. From these findings the conclusion was drawn that an interference with the coreceptor binding site is unlikely [22]. The cultivation of HIV-1_{NL4-3} in the presence of PEGI yielded in a resistant virus after several months of passaging. The presence of protein-binding agents like antibodies or peptides provokes the developing of specific mutations that will finally result in viral resistance. For instance, epitope specific mutation could be shown for glycan targeting antibodies and enfuvirtide in clinical trials [137, 138]. The PEGI-resistant HIV-1 obtained in this study showed five mutations within gp120 that differed distinctly from those detected in parallel cultivated controls. A useful benchmark for the evaluation of these mutations is the frequency of occurrence in the *Los Alamos HIV-1 sequence database* [139]. K33N and F175L can be commonly found in available sequences in the HIV-1 sequence database but V120E, S128N and A327T are notably rare (NL4-3 numbering). Although an asparagine was introduced, the consensus sequence (N-X-S/T) required for an N-glycosylation was not present due to the missing vicinal serine or threonine [140]. For structural analysis a variety of published structures are available, but only a few are displaying the entire variable region of gp120. The analysis of a prefusion trimeric ENV crystal structure led to the following observations: the V3 loop is neighbored by the V1, V2 complex and mutation sites (AA120 and AA175) are in close proximity to the V3 loop tip. The V3 loop base with mutation site AA 327 is shielded by several beta strands. The last mutation (F175L), which recovered the viral fitness and enabled resistance against inhibition by PEGI was located in the C1 region (Figure 57).

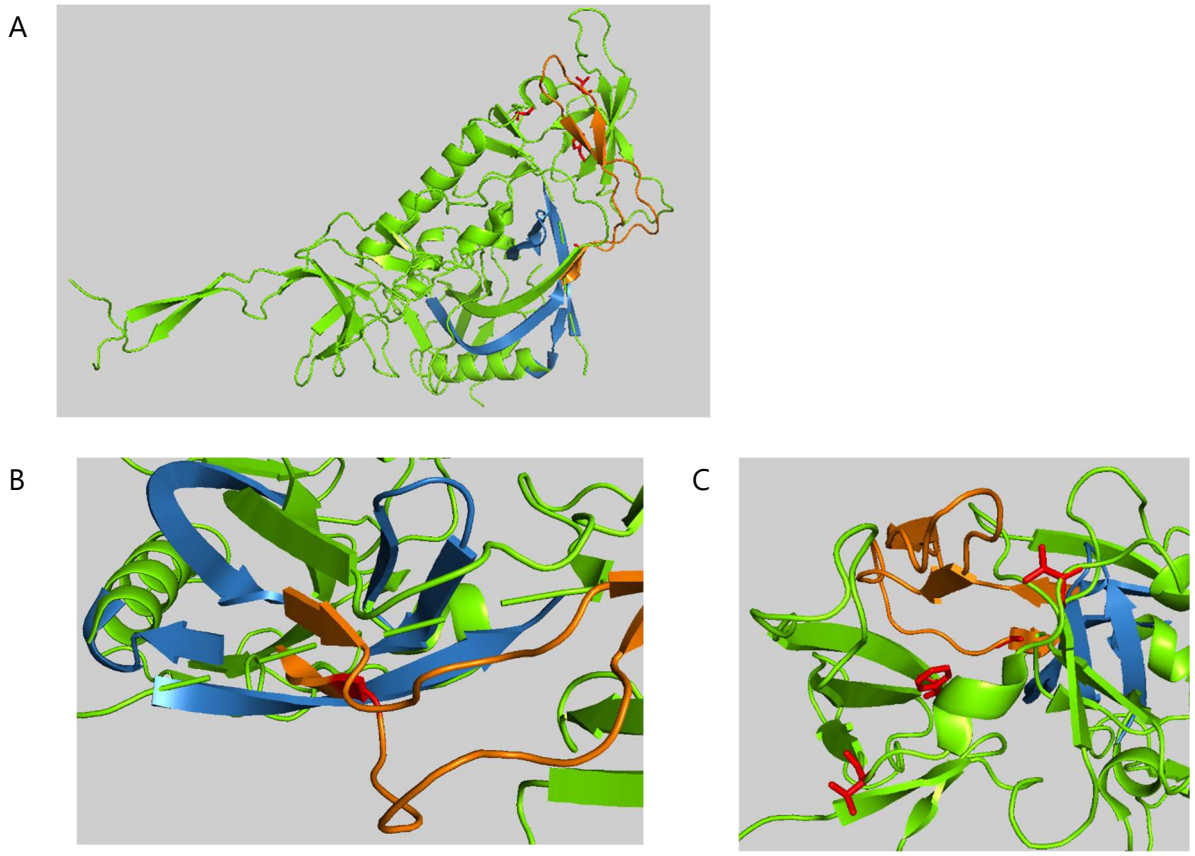
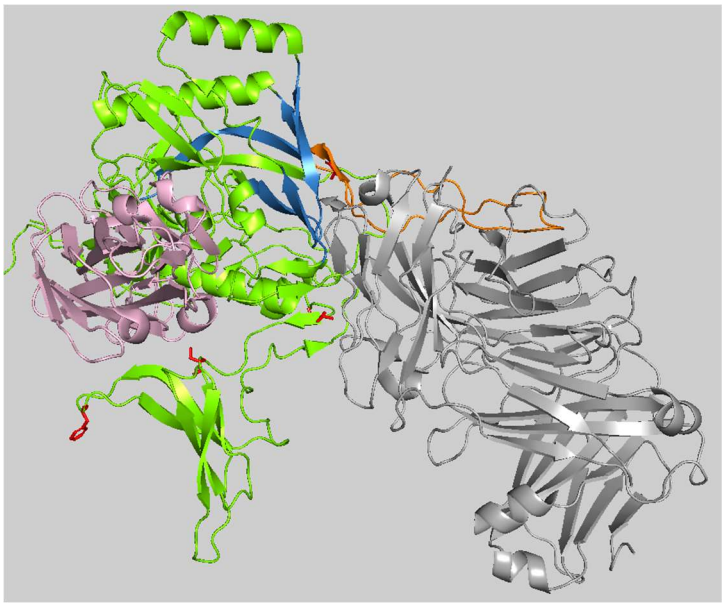


Figure 57: Prefusion trimeric ENV crystal structure section A: Section showing a gp120 monomer (green) with highlighted V3 loop (orange), C-terminal V3 loop base neighboring beta strands (blue) and mutation sites of PEGI resistant NL4-3 (red). B: Section of the crystal structure showing the V3 loop stem of gp120 with highlighted V3 loop (orange) C-terminal V3 loop base neighboring beta strands (blue) and A327 (red). C: Section of the crystal structure showing the V3 loop tip of gp120 with highlighted V3 loop (orange) C-terminal V3 loop base neighboring beta strands (blue) and mutation sites of PEGI resistant NL4-3 (red). pdb: 4TVP.

These four mutations cluster at the interface of V3 with the junction of V1 stem and C1 region in this non-bound ENV conformation. Especially, mutation site 175 and 120 are closely neighboring the tip of the V3 loop. This tip section of the V3 loop, due to its exposed accessibility, was identified as the epitope of a variety of neutralizing and non-neutralizing antibodies which were elicited during HIV-1 infection in humans [121, 141, 142]. The mutation sites 120, 128 and 175 are referred to as the mutation cluster in the following. The comparison of the conformational orientation in the mutation cluster of the prefusion crystal structure with a

recently published model of the CD4-bound ENV conformation in complex with mAb 17b showed that CD4 binding has a distinct effect on the orientation of the mutation cluster. CD4 binding causes profound conformation changes of the variable loops of gp120. The loops are prolapsed from the gp120 core with the V3 loop being orientated towards the interaction site with mAb 17b. 17b is an antibody that targets the coreceptor binding site, which is only accessible after CD4 binding [143]. As a result, the mutation site, in the V1-V2-loop region, obtained by long-term cultivation with PEGI, loses its neighboring contact site with the V3 loop. The beta strands shielding the V3 loop stem remain in the CD4 and 17b-bound conformation (Figure 58).

A



B

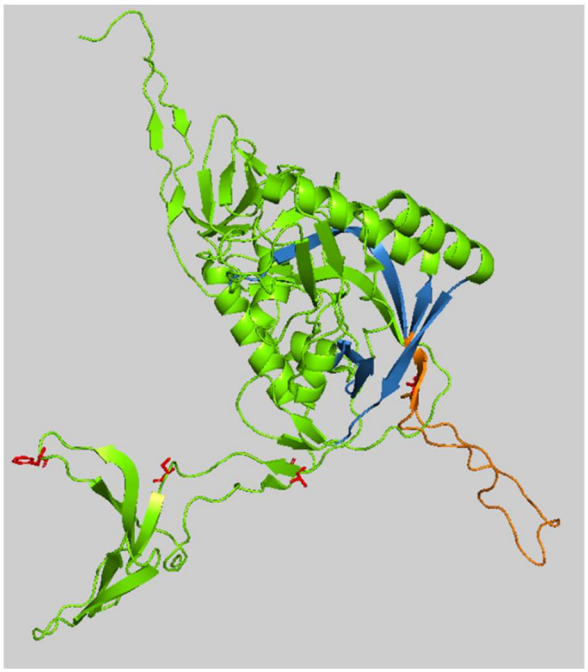


Figure 58: HIV-1 ENV spike model of a CD4-bound ENV conformation A: Section showing a gp120 monomer (green) bound to CD4 (pink) and 17b (grey) with highlighted V3 loop (orange) and mutation sites of PEGI resistant NL4-3 (red). B: Section showing only the gp120 monomer (green) highlighted V3 loop (orange) and mutation sites of PEGI resistant NL4-3 (red). pdb: 3J70.

A correlation between the mutation sites, obtained by long-term incubation and PEGIs ability to inhibit the binding of V3 loop antibodies in competitive assays can be postulated. The strongest ability to interfere with an antibody-protein binding was found for the V3 loop

antibodies, especially for F425 B4e8. CD4 binding site antibodies were susceptible to inhibition by PEGI as well. Although all tested V3 loop antibodies target the tip of the loop, the mode of interaction as well as the HIV-1 neutralizing capacity are different. The structural basis for these differences is evaluated by the available crystal structures of F425 B4e8 and 447-52D. The epitopes of both antibodies share the key residues, especially R315, but differ in the adopted V3 loop conformation and the main chain involvement [144, 145]. The key residue R315 is part of the GPGR motif at the tip of V3 which is in close proximity to mutation site AA120 in prefusion conformation (Figure 59) [121]. The mutation V120E also introduces a new possible site for salt bridges, which could either stabilize or destabilize a specific protein conformation [146]. The close distance in the prefusion crystal structure is indicating that an ionic bound between position 120 and positive charged AA at the very tip of the V3 loop is theoretically possible.

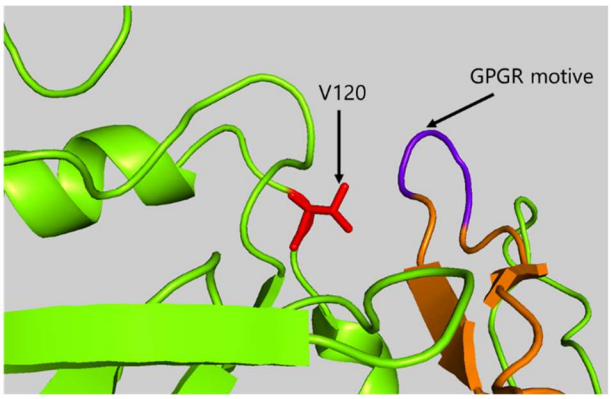


Figure 59: ENV crystal structure section showing position 120 and the tip of V3 loop of gp120. Highlighted in red = position 120, orange = V3 loop, in violet = GPGR tip motif of V3, pdb: 4TVP.

The hypothesis that at least a part of the PEGI binding site is directly located at or near the tip of the V3 loop is supported by the long-term incubation experiment with HIV-1_{NL4-3} as well as the results obtained in V3 loop antibody-$gp120_{HXBc2}$ competitive assay. In addition, the PEGIs ability to bind the V3 loop was proven in binding assay and in the context in HIV-1 neutralization assay. It can be speculated that binding of PEGI prevents either C1-V1-junction or V3 loop to adopt the necessary conformation crucial for further ENV rearrangement during host cell entry. The conformation of the V3 loop is known to be an important requirement for coreceptor binding and hence for the whole entry process [147, 148]. The base of the V3 loop

could be excluded as a possible epitope for PEGI because it is extensively shielded by beta strands and glycosylation in both conformations, prefusion and CD4-bound. In addition, a C-terminally neighboring glycan motif is the known epitope for 2G12, which was not susceptible to inhibition by PEGI in mAB-gp120$_{HXBc2}$ competitive assay. The inhibitory activity of PEGI was not exclusively selective for V3 loop targeting antibodies. CD4-binding site antibodies binding to gp120$_{HXBc2}$ was also inhibited by PEGI, but the proposed PEGI binding region is not overlapping with the epitope of the CD4 binding site antibodies [47, 48, 51]. A binding site with high similarity to the CD4 binding site is not considered to be a probable binding site for PEGI based on the mutation pattern of the PEGI resistant virus. Moreover, detected mutations were not correlating with resistant strain data available for b12, although position 120 and 128 are neighboring a part of the discontinuous CD4 binding site [149, 150]. Interestingly, binding of CD4 switches the orientation of the V1-C1 junction and maybe a mutation in the C1 region can influence conformational changes needed for the binding of CD4 (Figure 57 and Figure 58). Direct interaction of PEGI with the V1-C1-junction by binding or competitive assays has not been demonstrated yet. The assumption can be made that the interaction of PEGI with CD4 binding site antibodies is more likely due to steric or conformational effects changing the antibody epitope accessibility because published data did not show a CD4-dependent activity of E2-protein derives peptides [22, 151]. Although several single mutations could be detected during the performed long-term incubation in the presence of PEGI, the incubated NL4-3 virus was still susceptible until the appearance of the last mutation. The virulence of these specific strains was recovered by a single additional mutation at position 175. The mutation of F to L in the tip region of V2 was mandatory for the resistance against PEGI. The long-term incubated virus improved its fitness by introducing this specific mutation. It can be hypothesized that the mutation is a direct consequence of simultaneous peptide cultivation based on the observation that F175L mutation appears to be a general determinant for elevated viral fitness among clade B viruses [152]. The loss of the phenylalanine aromatic moiety could cause a rearrangement of the V2-V3 interaction side, which effectively shields PEGIs binding site and can be considered a possible mechanism for viral resistance. In particular, the introduction of leucine at the very tip of V2 in gp120 (position 175) with V120E and S128N mutations (NL4-3 numbering) can be found in a virulent clinical isolate (Genbank Accession number KY324517.1) [137]. The mutation pattern found in a clinical strain emphasized that the position pattern is not a coincidence but

interdependent and that the functional role of gp120 can be maintained regardless of the mutation of the conserved position 120 and 128. It can be concluded that the interaction with gp120 is a molecular determinant for HIV-1 inhibitory activity based on the ability of PEGI to effectively compete with gp120 targeting antibodies, its binding to monomeric gp120 and the peptide dependent mutations of the laboratory strain NL4-3 during long-term incubation. More details regarding the interaction site of PEGI were discovered attributing its HIV-1 inhibitory potential, especially its V3 loop binding, thereby either directly or indirectly destabilizing crucial gp120 conformation.

5.1.3 HIV-1 Inhibition by PEGI – a Clade Dependent Mechanism?

HIV-1 replication inhibitors can be tested with several cell-based in vitro systems. For this purpose, a variety of cell lines is available. A well-established HIV-1 neutralization assay uses CEMx174 M7R5 cells, which enables the testing of R5- and X4-tropic isolates in the same system. PEGI demonstrated that its HIV-1 inhibitory activity was not limited to clade B laboratory strains, as isolates from nearly all tested clades were sensitive to inhibition. Of the tested HIV-1 strain panel only isolates from CRF02 AG together with two isolates from clade C and F were unaffected. The resistant strains can be categorized by their coreceptor tropism. So far, only R-tropic strains were resistant against inhibition by PEGI. A common mechanism for CRF02-AG strains might be possible with both tested strains exhibiting resistance. The basis for the CRF02-AG sequence is a mosaic of clade A and clade G sequence sections. This is accompanied by a superordinate sequence independent replicate fitness compared to the derived clades. The gain in viral fitness is considered to be the main reason for speed and breadth of distribution after evolving this circulating recombinant form [153]. It is plausible that enhanced viral fitness might also be the reason for resistance against neutralization by PEGI.

The proposed hypothesis for the inhibition mechanism is the direct or indirect interaction of PEGI with the V3 loop of gp120. The V3 loop of AG01 varies distinctly from clade B and lacking the parts of “PxR” tip motif. As a result, antibodies like 447-52D were less active and V3 loop antibodies elicited against AG01 strains were not cross-reacting with clade B strains [154, 155]. Due to their viral fitness the CRF AG01 representatives of the HIV-1 subtype panel have a

special status and therefore their resistance was elucidated in detail. The comparison of the aligned sequences with NL4-3 (used as standard reference for clade B) exhibits prominent differences in the length of either V1 or V2 loop. The V1 loop of AG02 is elongated by 6 AA whereas a truncated V2 loop missing 7 AA and an elongated V2 loop (+ 9 AA) is observed for AG01 compared to NL4-3 (Figure 60 and 7.2, Figure 74). These changes in the length of either V1 or V2 loop could have an impact on the conformation of the variable regions of gp120 and therefore prevent the binding of PEGI to its gp120 located binding site.

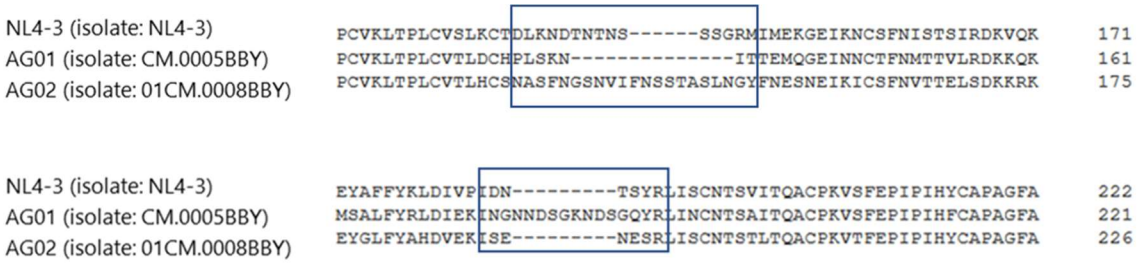


Figure 60: Section of the sequence alignment of strain NL4-3, AG01 and AG02. Blue boxes are highlighting sequence elongations of AG01 or AG02 compared to NL4-3 Alignment created with *Clustal Omega* [119].

In contrary to the resistance of AG01 and AG02, the tested strains from clade A and G were sensitive to inhibition by PEGI. The potency of PEGI in the tested HIV-1 panel ranged from a nanomolar IC_{50} to resistance at 60 µM, depending on the specific isolate. The assumption could be made that the resistance is caused by a PEGI sequence independent factor. In the following it will be discussed whether a PEGI sequence dependent resistance is plausible for C2 and F1. For this purpose, ENV sequences of C2, F1 and NL4-3 were aligned together with two representatives of highly susceptible strains (G ARP and SF33). The focus of the sequence analysis was laid on variable loops 1-3 (V1, V2 and V3) including their stem regions where PEGI induced mutations were detected during long-term experiments with NL4-3. Mutation K33N and F175L (NL4-3 numbering) observed after long-term incubation in the presence of PEGI can be also found in resistant and non-resistant strains and are therefore not expected to be essential for the strain specific resistance of C2 and F1. The chain elongation with introduction of an additional glycosylation site within the V2 loop of F1, compared to clade B strains and F2, is a possible sequence-based reason for resistance against PEGI (Figure 61). This additional

glycosylation site could cause more extensive shielding of the protein surface and maybe limit the accessibility of the PEGI binding site. The analysis of the sequence alignment did not lead to a potential sequence-based reason for C2 resistance. Testing more isolates of clade C would be useful to investigate the inhibition mechanism of PEGI in clade C HIV-1 isolates. A binding site independent weaker neutralization of clade C compared to clade A and B was described for V3 loop antibodies. Clade C strains are less susceptible because they are lacking the flexible β-hairpin (AA 307-309, HXBc2 numbering) [156]. It is possible that this binding site independent effect could also apply for the resistance of isolate C2 against PEGI. This would also support the V3-loop binding site hypothesis for PEGIs inhibition mechanism. The testing of more than two representative of clade C and F is needed to clarify in detail if a clade dependent resistance mechanism exists. The mechanism can be either sequence-based or due to other virus properties like ENV glycosylation pattern or general higher viral fitness against neutralization by ENV targeting entry inhibitors like PEGI. The resistance of AG01 and AG02 and their distinct elongation in the variable loops, especially loop 2, indicates a higher likelihood for a strain specific mechanism.

Figure 61: Section of the sequence alignment of strain NL4-3, G ARP, C2, F1 and SF33. Blue boxes = mutation site of the long-term incubation experiment conducted with PEGI and NL4-3, orange box = elongation of F1 compared to the other aligned sequences, * = no difference, . = one strain different AA, : two strains same AA. Alignment created with Clustal Omega [119].

The available results of PEGI tested in HIV-1 strain panel is indicating that PEGI resistance can be driven by the coreceptor selectivity of the virus. The HIV-1 strain panel tested in HIV-1 neutralization assays also emphasized the cell-line dependency concerning susceptibility and virulence of the specific HIV-1 isolate. Previously published infection assay results using a TZM-bl cell line showed PEGI resistance of isolates C1, F1, G/H, as well as three additional tested R5-tropic HIV-1 isolate belonging to clade B,D and F [85]. HIV-1 isolates C1, F1 and G/H

were susceptible to inhibition by PEGI in the HIV-1 neutralization assay using the CEMx174 M7R5 cell line, presented in this work. Despite the variation between these two assays, only R5-tropic (including one dual tropic isolate) were resistant against PEGI. The capacity of PEGI to inhibit infection of HIV-1 subtypes is not determined by the clade. It can be assumed that CF02 AG strains have a specific resistance mechanism whereas the resistance of the other strains is most likely to be directed by their tropism or isolate individual features but not by their subtype. Tropism dependent resistance is described for gp120 targeting broad neutralizing antibodies PG9 and PG16 which target a quaternary epitope within the V1-V2 loop region. The study showed a significant increased resistance of X4-tropic isolates [157]. Certainly, the number of isolate samples needs to be increased to further elucidate if the resistance selectivity of R5-tropic is significant for the HIV-1 inhibitory peptide PEGI.

Another approach was to analyze the resistance pattern of the HIV-1 strain panel based on the regional predominance of HIV-1 and HPgV-1. The peptide sequence of PEGI is derived from a strain belonging to HPgV-1 virus subtype 2. This is the predominant subtype for America and Europe (c.f.1.3.1). In the same way HIV-1 subtypes of the tested HIV-1 strain panel can also be sorted by their distribution areas. The distribution pattern is clade dependent; some are only occurring in Africa whereas others are globally spread as well as being the predominant strain for whole (sub)continents (Figure 62).

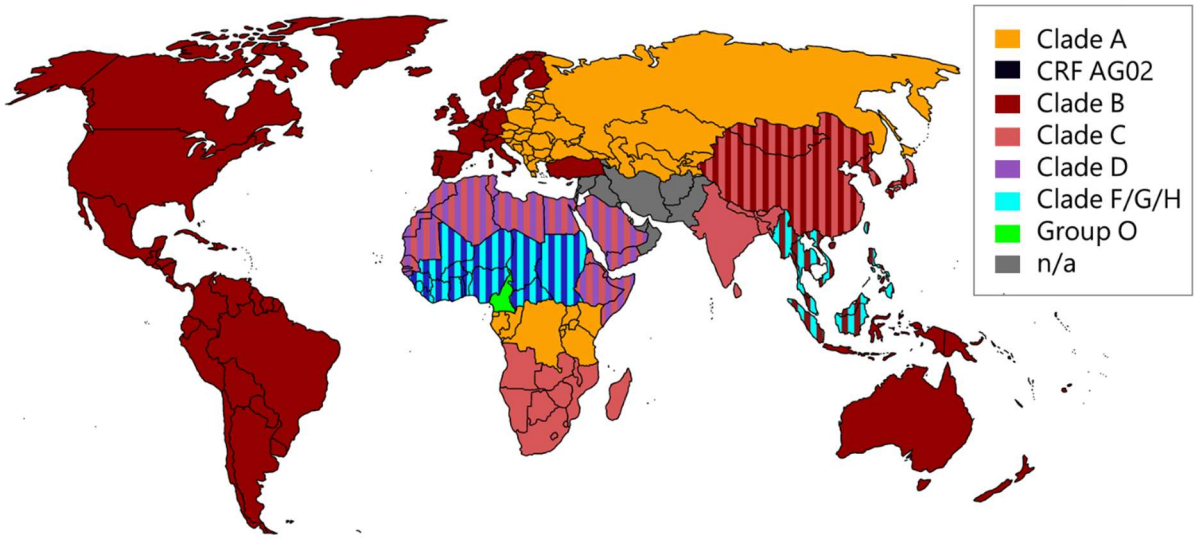


Figure 62: World map with highlighted predominant distribution of clades and group O which were present in the tested HIV-1 strain panel [58, 158].

Three out of four resistant strains from the HIV-1 strain panel have their origin in Central Africa where HPgV-1 subtype 2 is not commonly distributed. In addition, all six tested clade B strains that share the same distribution pattern as HPgV-1 subtype 2 were inhibited by PEGI. The distribution analysis of HIV-1 and HPgV-1 subtypes suggested a correlation between susceptibility to inhibition by PEGI and HPgV-1 subtype 2 occurrence. Consequently, the vulnerability against inhibition by PEGI is expected to be more probable for HIV-1 strains that have a common distribution with HPgV-1 subtype 2 [114]. This implies that the viral interference based on the interaction of E2 protein derived peptides with HIV-1 depends on a co-infection likelihood. It can be speculated that this dependency is a reason for the variability of available reports about the beneficial outcome of the co-infection of HIV-1 infected individuals [75]. Taking together these considerations, the HIV-1 inhibition by PEGI is mainly directed by two aspects: tropism and common distribution area with HPgV-1 subtype 2.

5.2 Using Mixture-based Libraries to Identify D-Hexapeptides which Inhibit VRC01 Binding to $gp120_{HXBc2}$

For the identification of new inhibitory peptides, mixture-based libraries have been screened for their ability to inhibit the $gp120_{HXBc2}$-VRC01 interaction. These peptide libraries belong to the group of positional scanning synthetic combinatorial libraries (PS-SCLs). An inhibitory active D-AA library was selected for further positional screening up to the level of deconvoluting individual peptides. Two peptides, called VRC01-ReH11 and VRC01-ReH12, were the two active peptides from the set of 24 synthesized and tested peptides. These peptides were identified by screening of a positional scanning mixture-based D-hexapeptide library (TPI 2338) and subsequent deconvolution of individual peptides. Both peptides are sharing a 5mer sequence overlap and only the C terminal amino acid is variable (Figure 63).

Figure 63: Most active peptide sequences derived from the positional scanning profile of TPI 2338 in $gp120_{HXBc2}$-VRC01 competitive assay.

It is noteworthy that the two-fold increase of inhibitory activity from the scaffold ranking sample to the individual peptide (VRC01-ReH12) was lower than expected (Table 39).

Table 39: Comparison of inhibitory activity of scaffold ranking sample, PS mixture position 1 with e as defined AA and VRC01-ReH12.

	All X sample	PS mixture position 1, Ac-exxxxx-NH_2	VRC01-ReH12
c [mg/ml]	0.63	0.31	0.18
AVG [% Inhibition]	57	51	69

Usually increasing activity while increasing the actual concentration of the active compound from initial mixture testing to the purified individual peptide is reported [92, 159]. The low activity increase detected for the active deconvoluted individual peptides is indicating that unselective hydrophobic interaction might contribute to the inhibitory activity. The peptide sequences show the potential of being prone for non-specific hydrophobic interaction due to the high proportion of hydrophobic AAs, up to 5 out of 6, within the sequence. In addition, the individual

active peptide failed to completely inhibit the $gp120_{HXBc2}$-VRC01 interaction and no convincing binding strength to either VRC01 or $gp120_{HXBc2}$ was detectable. This also indicates that unspecific hydrophobic interactions could play a role for the ability to inhibit the interaction (cf. 4.2.3). In the used competitive assay system specific inhibition can be caused by binding of the competing peptide to either $gp120_{HXBc2}$ or VRC01. The likelihood of unspecific inhibition can be therefore determined by testing the ability to bind to $gp120_{HXBc2}$ and to VRC01. Thus, only moderate inhibitory activity of the active peptide and no significant binding to both interaction partners is emphasizing that an unspecific inhibition could be possible. To further assess low activity of the active peptides, an additional evaluation of the positional scanning profile was performed. The FOX profile, especially for positions 1-3, was not very differentiated (c.f. 4.2.2). In general, the positional scanning profile was challenging and a deconvolution was only made possible by introducing the FOX based analysis. This implies two considerable aspects: the contribution of more than one individual sequence to the overall library activity and the likelihood of not selecting the best AA per position.

Positional scanning profiles with a low level of differentiation can be an indication that more than one peptide is active. Nevertheless, it can also indicate that only low inhibitors are present in the screened library. It is likely that many similar active peptides could have synergistic effects in the used $gp120_{HXBc2}$-VRC01 competitive assay approach. The binding of the short hexapeptides could either cover different parts of the discontinuous VRC01 epitope or simply bind to either the antibody or the protein. The VRC01 epitope within gp120 is a so-called discontinuous epitope, which means that it is not a linear section but covers different parts of gp120 (c.f. 1.2.3).

Selecting only a small number of peptides (24 in this case) was bearing the risk of not selecting the most active sequence. This risk is elevated when the scanning profile has a lower level of differentiation. To further address the question, whether a more active peptide can be found in this library, the testing of a bigger set of deconvoluted peptides will be needed. Nevertheless, the low level of differentiation of the positional scanning profiles is also indicating that the overall inhibitory activity of the individual peptides of TPI 2338 might be in the same range that was found for VRC01-ReH11 and -ReH12.

Furthermore, the inhibitory activity of VRC01-ReH11 and -ReH12 was not limited to the $gp120_{HXBc2}$-VRC01 interaction. Both peptides, identified *via* screening in the $gp120_{HXBc2}$-VRC01 competitive assay, were also able to inhibit the b12-$gp120_{HXBc2}$-interaction. The activity in $gp120_{HXBc2}$-b12 competitive assay can be explained by the results of the TPI 2338 positional scanning samples. In addition to the present VRC01 results, TPI 2338 positional scanning samples were also tested in $gp120_{HXBc2}$-b12 competitive assay. The direct comparison of both obtained FOX matrices revealed an interesting correlation. The AAs per position corresponding to VRC01-ReH11 and -ReH12 also showed a median FOX higher than 1 in the b12 positional scanning profile. In particular, these defined AAs were in the upper third of the profile of each position (Table 40 and chapter 7.2, Figure 75). Consequently, the inhibitory activity in $gp120_{HXBc2}$-b12 competitive assay was not a coincidence but rather predictable by the obtained positional scanning profiles. Because both antibodies belong to the same class of CD4-binding site antibodies with partly overlapping epitopes, a sequence overlap of active peptides can be expected [29].

Table 40: FOX matrices of positional scanning of TPI 2338 with $gp120_{HXBc2}$-VRC01 and $gp120_{HXBc2}$-b12 competitive assay, color code per position: lowest value = red, quantile = yellow, highest value = green.

$gp120_{HXBc2}$-VRC01 competitive assay

	FOX Matrix					
	P1	P2	P3	P4	P5	P6
a	0.94	0.87	0.92	0.81	1.09	0.91
c	0.87	0.90	1.08	1.12	0.77	1.07
d	0.92	0.56	0.94	0.78	0.96	1.01
e	1.24	1.00	0.92	0.87	1.19	0.97
f	1.14	1.22	1.06	1.06	0.92	0.83
g	1.01	0.68	0.80	0.67	0.76	0.86
h	0.99	0.89	0.76	0.88	1.01	0.85
i	1.22	1.16	1.27	1.03	1.18	1.06
k	0.99	1.09	1.00	0.86	0.99	0.89
l	1.17	1.16	1.20	1.06	0.99	1.06
m	1.07	1.07	1.08	1.11	0.91	0.96
n	0.77	0.67	0.77	0.74	0.80	0.83
p	0.99	0.54	0.81	0.72	0.58	0.77
q	0.98	0.80	0.76	0.97	1.01	0.96
r	1.00	1.03	0.78	0.97	0.86	0.95
s	1.13	1.06	0.95	0.73	1.06	1.01
t	1.17	1.13	1.04	0.70	1.07	0.98
v	1.25	1.01	1.25	1.09	1.19	1.09
w	1.10	0.84	1.16	1.09	0.99	1.00
y	1.18	1.18	1.13	1.16	1.08	1.06

$gp120_{HXBc2}$-b12 competitive assay

	FOX Matrix					
	P1	P2	P3	P4	P5	P6
a	0.79	1.00	0.74	0.54	0.82	0.93
c	0.66	0.84	1.06	1.21	0.63	0.94
d	0.51	0.45	0.33	0.93	0.50	0.88
e	1.11	0.67	0.71	1.00	1.19	0.80
f	1.07	1.22	1.13	1.13	0.92	0.82
g	0.88	0.49	0.54	0.70	0.68	0.86
h	0.85	0.63	0.67	0.41	0.68	0.85
i	1.35	1.28	1.32	1.01	1.09	1.14
k	1.05	1.32	0.87	0.96	1.02	0.80
l	1.12	1.23	1.19	1.11	1.05	1.13
m	0.96	1.03	1.05	1.08	0.68	0.88
n	0.61	0.60	0.48	1.03	0.70	0.79
p	0.68	0.44	0.64	0.93	0.72	0.48
q	0.82	0.72	0.47	0.88	0.78	0.90
r	0.99	1.22	0.86	1.26	0.94	1.20
s	1.03	0.90	0.82	0.54	0.96	0.86
t	1.36	1.02	1.14	0.85	0.84	0.87
v	1.27	1.16	1.26	1.18	1.18	1.09
w	1.15	0.82	0.97	0.92	1.07	0.90
y	1.19	1.14	0.89	1.16	0.89	1.05

Despite the fact that the inhibitory activity of the active peptides was lower than expected, some degree of specificity could be shown. It should be noted that the tested peptides are completely composed of D-AAs. The L-AA variants of the active peptides (VRC01-ReH11 and -ReH12) were unable to inhibit the $gp120_{HXBc2}$-VRC01 interaction which indicated that the inhibition of the active peptides is D-AA specific. However, the binding to either $gp120_{HXBc2}$ or VRC01 could not explicitly shown for the active peptides.

Non-binding but inhibitory peptides could have two explanations: Inhibition can be caused by an unspecific non-binding related mechanism or by a specific binding mechanism that is not

accessible in the used direct binding assay. The limitation of an ELISA-based direct binding assay is that ligand-binding caused conformational changes of the $gp120_{HXBc2}$-VRC01 interaction cannot be displayed because only one interaction partner is simply present. A second factor that needs to be considered is a biased concentration range. It could either be too low or too high. To clarify if the active peptides are really non-binding to either $gp120_{HXBc2}$ or VRC01 in direct binding assays, different setups should be tested. Two active peptides were identified that inhibit the $gp120_{HXBc2}$-VRC01 interaction. Sequence-dependent activity was confirmed by SAR (structure-activity relationship) controls and the activity was D-AA specific, but the overall inhibitory activity was lower than expected.

The active peptides can be a starting point for further optimization of the peptide potency. The combinatorial library approach offers the opportunity to study the effect of the extension of the current sequence [92]. A suitable approach would be to create a positional scanning decapeptide with six already defined positions. The defined position would correspond to the previous identified active hexapeptide (VRC01-ReH12). Three different arrangements of the newly introduced position 7-10 could be investigated (Figure 64).

Ac-xxefvwvyxx-NH_2 Ac-xxxxefvwvy-NH_2 Ac-efvwvyxxxx-NH_2

Figure 64: three different possible arrangements of a positional scanning decapeptide library based on VRC01-ReH12, x = undefined.

Several approaches were pursued to identify new HIV-1 entry inhibitors in the last years. For D-AA peptides, two targets are already reported: the coreceptor CXCR4 and the hydrophobic binding pocket of gp41. A D-AA decapeptide (Ac-rrmyrriyrr-NH_2) was characterized as a potent CXCR4-antagonist which was obtained by a combinatorial library screening approach [160]. Moreover, a polyarginine (Ac-rrrrrrrrr-NH_2) was found to be an active HIV-1 inhibitor [161]. The D-AA containing 20mer named DV1 was found to be inhibiting X4-tropic strains in HIV-1 replication assay. This peptide is representing a section of the N-terminus of the viral chemokine called vMIP-2 which is a known ligand of the CXCR4 receptor [162]. The second already identified target for D-AA peptides in the context of HIV-1 entry inhibition is located in the N heptad repeat of gp41. Cyclic D-AA peptides, monomeric as well as trimeric variants, targeting this specific region were reported as inhibitors. The first active peptide belonging to

this group was discovered by phage display [163, 164]. The inhibition of the interaction of CD4 binding site antibodies with gp120 by D-AA peptides was a new finding. Peptides totally composed of D-AA offer enhanced proteolytic stability because of their opposite chirality to the proteinogenic L-AAs that can be processed by cells [165]. Therefore, D-peptides are in the focus for the development of therapeutically administered drugs.

So far, only L-configurated peptides were reported as epitope mimics of VRC01 and b12 [166-168]. For b12 a paratope mimetic is also reported [10]. Apart from being composed of L-AAs, the comparison of the reported peptide sequences with the newly identified inhibitory D-AA peptides revealed two additional different sequence characteristics. The reported sequences ranged from a cyclic decapeptide up to a linear peptide composed of 45 amino acids [10, 166-168]. The D-AA hexapeptides which were identified in this work are notably shorter. A minimum chain length appeared to be mandatory to generate highly active epitope mimics. It is likely that the chain length of six of the VRC01-ReH peptide could be too short to achieve high inhibition. This especially needed to be considered for the CD4-binding site antibody-gp120 interaction, which is based on the binding to a discontinuous epitope covering different sequence sections of gp120. Consequently, the proposed extension of the sequence length could be beneficial to increase activity.

Except for the linear 10mer called E1 described by *Chikaev et al.*, the peptides were rather conformationally constrained than flexible: the C1 peptide described as an VRC01 epitope mimic is cyclized, the b12 mimotope, which is a homodimer, linked *via* a disulfide bridge, and the b12 paratope mimetic sequence contains an intermolecular disulfide and a lactam bridge [10, 166-168]. Reducing the sequential flexibility appeared to be beneficial for interfering the CD4-binding site antibody-gp120 interaction by peptides. It has been shown that increasing the rigidity of the peptide has a positive effect on the overall bioactivity [169]. The identified hexapeptides VRC01-ReH11 and -ReH12 are lacking conformationally restriction *via* linkage or cyclization. To further optimize the inhibitory activity, cyclized variants of the identified peptide sequences could be tested. Moreover, several library samples showed promising activity inhibiting the $gp120_{HXBc2}$-VRC01 interaction in the initially performed scaffold ranking screening (c.f. 4.2.1). The further testing of the most active cyclic scaffold library TPI 2321 would be a promising candidate to identify an inhibitory peptide with a higher potency than already identified via screening the D-AA hexapeptide library (TPI 2338).

The identified active hexapeptides completely composed of D-AA demonstrated that the established antibody-protein competitive assays were successfully applied to detect compounds that inhibit the designated interaction. The D-AA hexapeptides VRC01-ReH11 and -ReH12 were identified by positional scanning analysis with the new FOX approach and demonstrated their D-AA selective inhibition of VRC01-$gp120_{HXBc2}$ interaction and sequence-selective inhibition of CD4-binding-site-$gp120_{HXBc2}$ interactions.

The first reported inhibitory peptides, composed of D-AAs, for the VRC01-$gp120_{HXBc2}$ interaction have proven the feasibility of using PS-SCLs in the context of HIV-1 ENV targeting antibody assays. Towards the goal of identifying highly-active peptides that inhibit the investigated antibody-protein interactions it is recommended to further test positional scanning sets of cyclic PS-SCLs, which showed inhibitory activity in the already performed scaffold ranking.

6 Summary

6.1 Summary

Peptides are versatile tools to investigate protein-protein interaction on the amino acid sequence level. The entry to the host cell of HIV-1 is a crucial step of the viral replication cycle. The investigation of inhibitory peptides was a promising starting point to gain further insight into the molecular mechanism of viral host cell entry. Moreover, the characterization of the structural activity of peptides as HIV-1 entry inhibitors contributed to a better understanding of the complex process of HIV-1 entry and its inhibition respectively. During this work two different approaches were pursued.

The first approach used a known inhibitory protein-derived peptide (PEGI) as a starting point. The PEGI peptide is derived from the E2 protein of HPgV-1, a non-pathogen human virus, which is a commonly occurring co-infection of HIV-1 infected individuals. Several epidemiological studies showed that a co-infection with HPgV-1 has a positive influence on the outcome parameters of the HIV-1 infection for example disease progression or mortality. The molecular determinants for this phenomenon called viral interference had been studied after the discovery of HPgV-1 in the late 1990s. It has been discovered that the envelope protein (E2), as well as an E2 protein-derived peptide are inhibitors of the HIV-1 infection in vitro. One of these peptides was the selected lead sequence for the PEGI peptides of this work. A range of PEGI sequence variants were synthesized and several prerequisite sequence properties for HIV-1 inhibitory activity were determined. The C-terminal negative charge, the hydrophobic core motif and the presence of three cysteines were crucial sequence characteristics for maintaining HIV-1 inhibitory activity. In addition, reduced cysteines instead of a connection via inter- or intramolecular disulfide or lactam bridges were found to be beneficial for HIV-1 inhibitory activity and thus, it can be speculated that redox-active cysteines are involved in the HIV-1 inhibition by PEGI. Moreover, truncation variants revealed a five-fold more potent HIV-1 peptide inhibitor (PEGI-D35), which is a promising candidate for follow up studies (Ac-LCDCPNGPWVWVPAFCQD-NH_2). The PEGI sequence represents a highly conserved motif within the E2 protein with only three flexible positions among the circulating HPgV-1 isolates. The relevance of the PEGI sequence for HPgV-1 isolates was proven by testing a panel of mainly

occurring variants of these E2 section and HPgV-1 strain variants, which were inhibitory active. The breadth of HIV-1 inhibition was determined with a HIV-1 strain panel. 4 of 20 tested HIV-1 isolates were resistant against inhibition by PEGI. However, inhibition with IC_{50} values up to the nanomolar range, depending on the PEGI sensitive HIV-1 isolate, was discovered. The inhibitory activity of PEGI was not limited to a single HIV-1 isolate and PEGI showed the potential of a cross-clade inhibitory peptide. Due to the fact that only R5-tropic HIV-1 isolates were resistant a preference for X4-tropic isolates is indicated. In addition, it could be shown that PEGI-sensitive HIV-1 isolates share the same distribution areas as the HPgV-1 subtype 2. The PEGI sequence was initially derived from the E2 protein of a subtype 2 HPgV-1 isolate. Nevertheless, PEGIs activity is not considered to be the only determinant for the observed complex phenomenon of viral interference between HIV-1 and HPgV-1. It can be concluded that PEGI is not a complete functional E2 protein mimic but still provided a mechanism for viral interference on the peptide-protein level. The successful cultivation of a PEGI resistant $HIV\text{-}1_{NL4\text{-}3}$ virus and subsequent sequencing of the viral genes discovered five single mutations within gp120, partly located in conserved sections of the protein. Especially the mutation of conserved positions suggested that an important binding site of PEGI is located within gp120.

Binding experiments were performed with a set of PEGI peptides to further support this hypothesis. The results outstandingly correlated with the results in HIV-1 neutralization assay. Based on this correlation, gp120 binding of PEGI is a plausible mechanism enabling HIV-1 inhibition. Moreover, it could be shown that PEGI inhibits the binding of gp120-targeting antibodies to their epitope. It was an especially potent inhibitor of the V3 loop antibody F425 B4e8 -gp120 interaction. The inhibition could be narrowed down to the epitope level in binding and HIV-1 neutralization experiments. After a profound analysis of all available results it could be concluded that the PEGIs HIV-1 inhibitory activity is based on binding to gp120, especially the V3 loop, and thereby either directly or indirectly destabilizing crucial gp120 conformation for a successful HIV-1 entry process. The HPgV-1 protein-derived peptide PEGI, including its variants, analysis of mode of action and neutralization breadth provided new important insights into to understand the viral interference of HIV-1 and HPgV-1 and HIV-1 inhibition on the protein level. To further support the hypothesized molecular mechanism of HIV-1 inhibition by PEGI, a follow up on the peptide level can be recommended which uses V3 loop sequences of resistant and susceptible HIV-1 isolates.

The second approach was the screening of synthetic combinatorial libraries in HIV-1 neutralizing antibody assays. The aim was to identify new peptide sequences that inhibit a designated antibody-protein interaction. 18 scaffold ranking samples of peptide libraries were tested in antibody-protein competitive assays. The following HIV-1 neutralizing antibodies were selected for screening: b12, VRC01, VRC03, three gp120 targeting antibodies, and F240, a gp41 targeting antibody. The screening results proved the suitability of the scaffold ranking to detect active peptide libraries in the field of HIV-1 neutralizing antibodies. A linear D-AA hexapeptide scaffold was selected as a promising candidate to further identify individual peptides which could inhibit the VRC01-$gp120_{HXBc2}$ interaction. Following the principles for a successful deconvolution, positional scanning samples of the D-AA hexapeptide library were tested. The evaluation of the positional scanning profiles was based on the newly introduced evaluation parameter called fold over X (FOX). FOX enabled the deconvolution of profiles with a moderate level of differentiation. The 24 peptides, selected during the deconvolution process, were synthesized and their ability to inhibit VRC01-$gp120_{HXBc2}$ interaction was tested. Two active hexapeptides were identified with a high sequence homology of five positions. Both peptides, VRC01-ReH12 (Ac-efvwvy-NH_2) and VRC01-ReH11 (Ac-efvwvl-NH_2), inhibited the VRC01-$gp120_{HXBc2}$ binding with moderate inhibitory activity (IC_{30} = 2 respectively 12 µM), whereas L-AA variants of the active peptides did not show inhibitory activity. The activity was not limited to the VRC01-$gp120_{HXBc2}$ interaction because both peptides also inhibited the b12- $gp120_{HXBc2}$ interaction. The identified peptides were the first reported D-AA hexapeptide competitors of HIV-1 neutralizing antibodies targeting the CD4-binding site. The testing of positional scanning samples of an active cyclic peptide library can be recommended for further studies to identify new active peptides.

6.2 Zusammenfassung

Erforschung von HIV-1-Wirtszell-Interaktionen mit Peptiden, die vom humanen Pegivirus-1 und von kombinatorischen Bibliotheken abgeleitet sind

Peptide sind vielseitig verwendbare Werkzeuge um Protein-Protein-Interaktionen auf der Aminosäuresequenzebene zu untersuchen. Der Eintritt von HIV-1 in die Wirtszelle ist ein zentraler Schritt innerhalb des viralen Replikationszyklus. Die Untersuchung inhibitorischer Peptide war ein vielversprechender Ausgangpunkt um weitere Einblicke in die molekularen Mechanismen des viralen Wirtszelleintritts zu gewinnen. Außerdem trug die Charakterisierung der strukturellen Aktivität von peptidischen HIV-1 Inhibitoren zu einem besseren Verständnis des komplexen HIV-1 Eintrittsprozesses sowie zu dessen Inhibition bei. Im Rahmen dieser Arbeit wurden dafür zwei verschiedene Ansätze verfolgt.

Der erste Ansatz nutzte als Ausgangspunkt ein bekanntes inhibitorisches Peptid (PEGI). Das PEGI Peptid ist von dem E2 Protein des nicht-pathogenen Humanvirus HPgV-1 abgeleitet, welches häufig als Co-Infektion von HIV-1 positiven Individuen auftritt. Verschiedene epidemiologische Studien zeigten, dass eine HPgV-1 Co-Infektion einen positiven Einfluss auf *Outcome*-Parameter der HIV-1 Infektion wie Mortalität und Krankheitsverlauf hat. Nach der Entdeckung von HPgV-1s in den späten 1990er Jahren wurden die molekularen Determinanten für dieses Phänomen, welches als virale Interferenz bezeichnet wird, studiert. Dabei wurde sowohl die Inhibition der HIV-1 Infektion *in vitro* durch das Hüllprotein (E2 Protein) als auch von E2-Protein-abgeleiteter Peptide entdeckt. Eines dieser Peptide wurde als Leitsequenz für die PEGI Peptide dieser Arbeit ausgewählt. Mit verschiedenen synthetisierten Varianten der PEGI Sequenz wurden grundlegende Sequenzeigenschaften für die inhibitorische Aktivität bestimmt. Die C-terminale negative Ladung, das hydrophobe Kernmotiv und das Vorhandensein von drei Cysteinen waren entscheidende Charakteristiken um die inhibitorische Aktivität zu erhalten. Zudem wurde gezeigt, dass anstatt eine Verbrückung über inter- oder intramolekular Disulfide oder Lactame reduzierte Cysteine vorteilhaft für die inhibitorische Aktivität waren. Somit kann über eine Beteiligung von redox-aktiven Cysteinen an der HIV-1 Inhibition von PEGI spekuliert werden. Außerdem zeigten Verkürzungsvarianten einen fünffach potenteren HIV-1 Peptidinhibitor, welcher ein vielversprechender Kandidat für nachfolgende Untersuchungen ist (Ac-LCDCPNGPWVWVPAFCQD-NH_2). Die PEGI-Sequenz stellt ein

hochkonserviertes Motiv des E2 Proteins mit nur drei variablen Positionen innerhalb zirkulierenden HPgV-1 Isolaten dar. Die Relevanz der PEGI Sequenz für HPgV-1 Isolate wurde mit der Testung eines Sets aus häufig auftretenden Varianten des E2 Proteinabschnitts, die inhibitorisch aktiv waren, belegt.

Die Breite der HIV-1 Inhibition wurde mit einer Auswahl von verschiedenen HIV-1 Stämmen bestimmt. 4 von 20 getesteten HIV-1 Isolaten waren gegen die Inhibition durch PEGI resistent. Jedoch wurden bei den PEGI-sensitiven HIV-1 Isolaten Inhibitionen mit IC50 Werten bis in den nanomolaren Bereich bestimmt. Die inhibitorische Aktivität von PEGI war nicht auf ein HIV-1 Isolat begrenzt und PEGI zeigte das Potential eines Clade-übergreifenden inhibitorischen Peptids. Eine Präferenz für X4-trope Isolate wurde angedeutet, da nur R5-trope HIV-1 Isolate resistent waren. Zudem konnte gezeigt werden, dass PEGI sensitive HIV-1 Isolate sich das gleiche Verteilungsgebiet mit dem HPgV-1 Subtyp 2 teilen. Die PEGI Sequenz war ursprünglich vom E2 Protein eines HPgV-1 Subtyp 2 Isolats abgeleitet. Trotzdem kann nicht angenommen werden, dass PEGI die einzige Determinante für das beobachtete komplexe Phänomen der viralen Interferenz von HIV-1 und HPgV-1 ist. Es konnte geschlussfolgert werden, dass PEGI kein vollständig funktionales Mimetikum des E2 Proteins ist. Trotzdem wurde ein Mechanismus für die virale Interferenz auf Peptid-Protein Ebene bestimmt. Die erfolgreiche Kultivierung eines PEGI resistenten HIV-1_{NL4-3}-Virus und anschließende Sequenzierung der viralen Gene stellte fünf Einzelmutationen innerhalb des gp120, teilweise in konservierten Bereichen des Proteins, fest. Im Speziellen die Mutationen in konservierten Bereichen ließen vermuten, dass eine wichtige PEGI Bindungsstelle im gp120 lokalisiert ist. Um diese Hypothese zu bestätigen, wurden Bindungsexperimente mit einer Auswahl an PEGI Peptiden durchgeführt. Die Ergebnisse korrelierten bemerkenswert mit den Ergebnissen der durchgeführten HIV-1 Neutralisationsversuche. Auf Grundlage dieser Korrelation ist die Binding von PEGI an gp120 ein plausibler Mechanismus, der die HIV-1 Inhibition ermöglicht. Außerdem konnte gezeigt werden, dass PEGI die Bindung von gp120-adressierender Antikörper an ihr Epitop inhibierte. PEGI war ein besonders potenter Inhibitor der V3 *Loop* Antikörper F425 B4e8-gp120 Interaktion. Die Inhibition konnte mit Bindungs- und HIV-1 Infektionsexperimenten auf die Epitope Ebene eingegrenzt werden. Nach einer fundierten Analyse aller verfügbaren Ergebnisse konnte geschlussfolgert werden, dass PEGIs inhibitorische Aktivität auf der Binding an gp120, im Speziellen an den V3 *Loop*, beruht und

somit entweder direkt oder indirekt entscheidende gp120 Konformationen für einen erfolgreichen HIV-1 Eintrittsprozess blockiert. Die Analyse der inhibitorischen Aktivität und Breite des HPgV-1 Protein abgeleiteten Peptids PEGI, inklusive seiner Varianten, lieferte neue wichtige Erkenntnisse für das Verständnis der viralen Interferenz von HIV-1 und HPgV-1. Um die Hypothese für den molekularen Mechanismus der HIV-1-Inhibition durch PEGI weiter zu unterstützen, kann eine Weiterverfolgung auf Peptidebene empfohlen werden, bei der V3-*Loop*-Sequenzen von resistenten und empfindlichen HIV-1-Isolaten verwendet werden.

Der zweite Ansatz war das Screening von kombinatorischen Peptidbibliotheken in HIV-1 neutralisierenden Antikörper Versuchen. Das Ziel war es neue Peptidsequenzen zu identifizieren, die eine bestimmte Antikörper-Protein Interaktion inhibieren. 18 *Scaffold Ranking* Proben von Peptidbibliotheken wurden in Antikörper-Protein Kompetitionsversuchen getestet. Die folgenden HIV-1 neutralisierenden Antikörper wurden für das Screening ausgewählt: b12, VRC01, VRC03, drei gp120 adressierende Antiköper und F240, ein gp41 adressierender Antikörper. Die Screening Ergebnisse bewiesen die Eignung des *Scaffold Rankings* für die Detektion von aktiven Peptidbibliotheken in den verwendeten HIV-1 Antikörper Versuchen. Ein lineares D-Aminosäure Hexapeptidgerüst wurde als vielversprechender Kandidat für die folgende Identifizierung einzelner aktiver Peptide, welche die VRC01-$gp120_{HXBc2}$ Interaktion inhibieren könnten, ausgewählt. Den Grundprinzipien für eine erfolgreiche Dekonvolution folgend, wurden Positionsscreening-Proben der D-Aminosäure Hexapeptidbibliothek getestet. Die Auswertung des Positionsscreening Profils beruhte auf dem neu eingeführten Auswertungsparameter, der *fold over X* (FOX) genannt wurde. FOX ermöglichte die Dekonvolution von Profilen mit moderaten Differenzierungsniveau. Die 24 Peptide, ausgewählt beim Dekonvolutionsprozess, wurden synthetisiert und auf ihre Fähigkeit die VRC01-$gp120_{HXBc2}$ zu inhibieren getestet. Zwei aktive Hexapeptide mit einer hohen Sequenzhomologie von fünf Positionen wurden identifiziert. Beide Peptide, VRC01-ReH12 (Ac-efvwvy-NH_2) und VRC01-ReH11 (Ac-efvwvl-NH_2), inhibierten die VRC01-$gp120_{HXBc2}$ mit moderater inhibitorischer Aktivität (IC_{30} = 2 bzw. 12 µM). L-Aminosäure Varianten der aktiven Peptide hingegen inhibierten nicht.

Darüber hinaus inhibierten die beiden Peptide nicht nur die VRC01-$gp120_{HXBc2}$ Interaktion, sondern auch die Interaktion von b12 mit $gp120_{HXBc2}$. Die identifizierten Peptide waren die ersten beschriebenen D-Aminosäuren Hexapeptidkompetitoren von HIV-1 neutralisierenden Antikörpern, die die CD4-Bindungsstelle adressieren.

Die Untersuchung von Positionsscreening-Proben einer aktiven zyklischen Peptidbibliothek kann für weitere Studien zur Identifizierung neuer aktiver Peptide empfohlen werden.

7 Appendix

7.1 Sequences and LC-MS data of the synthesized peptides

All sequences of the peptides presented in this work are shown in Table 41. The table is followed by the LC-MS data of the listed peptides. The methods of the LC-MS analysis can be found in chapter 3.2.4.

Table 41: List of synthesized peptides.

Name	Sequence	Molecular mass [g/mol]
PEGI	Ac-LCDCPNGPWVWVPAFCQAVG-OH	2202
PEGI-Biotin	Bio-XLCDCPNGPWVWVPAFCQAVG-OH	2501
PEGI-D2	Bio-XLCDCPNGPWVWVPAFCQAVG-NH_2	2499
PEGI-D3	Ac-LCDCPNGPWVWVPAFCQAVG-NH_2	2204
PEGI-D4	Bio-XLSDSPNGPWVWVPAFSQAVG-OH	2452
PEGI-D5	Ac-LSDSPNGPWVWVPAFSQAVG-OH	2156
PEGI-D6	Bio-XLSDSPNGPWVWVPAFSQAVG-NH_2	2451
PEGI-D7	Ac-LSDSPNGPWVWVPAFSQAVG-NH_2	2156
PEGI-D8	Bio-XLCDSPNGPWVWVPAFCQAVG-OH	2485
PEGI-D9	Bio-XLSDCPNGPWVWVPAFCQAVG-OH	2485
PEGI-D10	Bio-XLCDCPNGPWVWVPAFSQAVG-OH	2485
PEGI-D11	Bio-XLCDSPNGPWVWVPAFSQAVG-OH	2469
PEGI-D12	Bio-XLSDCPNGPWVWVPAFSQAVG-OH	2469
PEGI-D13	Bio-XLSDSPNGPWVWVPAFCQAVG-OH	2469
PEGI-D14	Ac-LCDSPNGPWVWVPAFCQAVG-OH	2189
PEGI-D15	Ac-LSDCPNGPWVWVPAFCQAVG-OH	2189
PEGI-D16	Ac-LCDCPNGPWVWVPAFSQAVG-OH	2189
PEGI-D19	Bio-XDCLCAAWVPQVPGPNCFGWV-OH	2501
PEGI-D20	Bio-XGCACDPQPFWVWVGLCPVAN-OH	2501
PEGI-D21	Bio-XIYEGCPGIAFYFWQAMNNLG-OH	2632
PEGI-D23	Ac-LCDCPNGPAVWVPAFCQAVG-OH	2089
PEGI-D24	Ac-LCDCPNGPWAWVPAFCQAVG-OH	2177
PEGI-D25	Ac-LCDCPNGPWVAVPAFCQAVG-OH	2089
PEGI-D26	Ac-LCDCPNGPWVWAPAFCQAVG-OH	2177
PEGI-D31	Ac-LCDCPNGPWVWVPAFCQAVD-NH_2	2262
PEGI-D32	Bio-XLCDCPNGPWVWVPAFCQAVD-NH_2	2441
PEGI-D33	Ac-L{Dap-DCPNGPWVWVPAFE}SQAVG-OH	2199
PEGI-D34	Bio-XLCDCPNGPWVWVPAFCQAD-NH_2	2458
PEGI-D35	Bio-XLCDCPNGPWVWVPAFCQD-NH_2	2387
PEGI-D36	Bio-XLCDCPNGPWVWVPAFCD-NH_2	2258

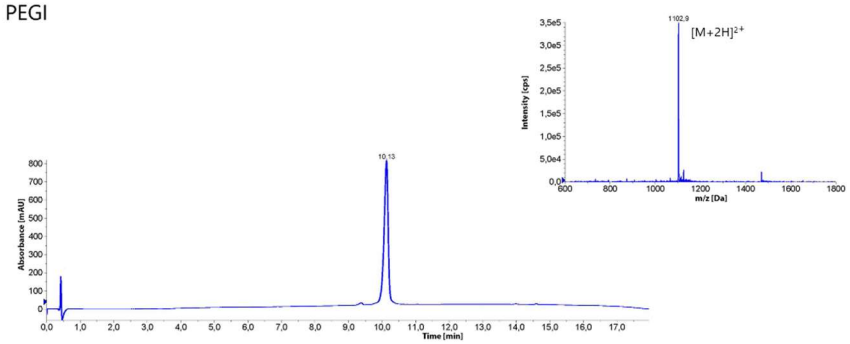
PEGI-D37	Bio-XCDCPNGPWVWVPAFCQAVG-OH	2387
PEGI-D43	Bio-XLMDCPNGPWVWVPAFCQAVG-OH	2528
PEGI-D44	Bio-XLCDMPNGPWVWVPAFCQAVG-OH	2528
PEGI-D45	Bio-XLCDCPNGPWVWVPAFMQAVG-OH	2528
PEGI-D46	Bio-XLMDMPNGPWVWVPAFMQAVG-OH	2587
PEGI-D47	Bio-XL{KDCPNGPWVWVPAFE}QAVG-OH	2532
PEGI-D48	Bio-XLCD{KPNGPWVWVPAFE}QAVG-OH	2532
PEGI-D49	Bio-XDCPNGPWVWVPAFCQD-NH_2	2169
PEGI-D50	Bio-XD[CPNGPWVWVPAFC]QD-NH_2	2167
PEGI-D51	Bio-XLCDCPNGPWVwVPAFCQAVG-OH	2501
PEGI-D52	Bio-XLCDCPNGPwVWVPAFCQAVG-OH	2501
PEGI-D53	Bio-XLCdCPNGPWVWVPAFCQAVG-OH	2501
PEGI-D54	Bio-XLCDCPNGPWVWLPAFCQAIG-OH	2531
PEGI-D55	Bio-XLCDCPNGPWVWMPAFCQAVG-OH	2534
PEGI-D56	Bio-XLCDCPNGPWVWQPAFCQAVG-OH	2531
PEGI-D57	Bio-XLCDCPNGPWVWSPAFCQAVG-OH	2490
PEGI-D58	Bio-XLCDCPNGPWVWHPAFCQAVG-OH	2540
PEGI-D59	Bio-XLCDCPNGPWVWVPALCQAVG-OH	2468
PEGI-D60	Bio-XDCPNGPWVWVPAFCQAVG-OH	2286
PEGI-D61	Bio-XLCDCPNGPWVWVPAFcQD-NH_2	2390
PEGI-D62	Bio-XLcDCPNGPWVWVPAFCQD-NH_2	2390
PEGI-D63	Bio-XLCDcPNGPWVWVPAFCQD-NH_2	2390
PEGI-D64	Bio-XL-Hcy-DCPNGPWVWVPAFCQD-NH_2	2402
PEGI-D65	Bio-X-LCD-Hcy-PNGPWVWVPAFCQD-NH_2	2402
PEGI-D66	Bio-X-LCDCPNGPWVWVPAF-Hcy-QD-NH_2	2402
PEGI-D67	Bio-XLCDCPNGPWVWVPAFCQAVGWGD-NH_2	2859
PEGI-D68	Bio-XLCDCPNGPWVWVPAFCQAVGWGDPIT-NH_2	3171
VRC01-crude-1	Ac-vfvyvl-NH2	780
VRC01-crude-2	Ac-vfvyvy-NH2	830
VRC01-crude-3	Ac-vfvyel-NH_2	810
VRC01-crude-4	Ac-vfvyey-NH_2	860
VRC01-crude-5	Ac-vfvmvl-NH_2	748
VRC01-crude-6	Ac-vfvmvy-NH_2	798
VRC01-crude-7	Ac-vfvmel-NH_2	778
VRC01-crude-8	Ac-vfvmey-NH_2	828
VRC01-crude-9	Ac-vfvwvl-NH_2	803
VRC01-crude-10	Ac-vfvwvy-NH_2	853
VRC01-crude-11	Ac-vfvwel-NH_2	833
VRC01-crude-12	Ac-vfvwey-NH_2	883
VRC01-crude-13	Ac-efvyvl-NH_2	810
VRC01-crude-14	Ac-efvyvy-NH_2	860
VRC01-crude-15	Ac-efvyel-NH_2	840
VRC01-crude-16	Ac-efvyey-NH_2	890
VRC01-crude-17	Ac-efvmvl-NH_2	778
VRC01-crude-18	Ac-efvmvy-NH_2	828
VRC01-crude-19	Ac-efvmel-NH_2	808
VRC01-crude-20	Ac-efvmey-NH_2	858
VRC01-crude-21	Ac-efvwvl-NH_2	833

VRC01-crude-22	Ac-efvwvy-NH_2	883
VRC01-crude-23	Ac-efvwel-NH_2	863
VRC01-crude-24	Ac-efvwey-NH2	913
VRC01-ReH2	Ac-vfvmel-NH_2	778
VRC01-ReH3	Ac-vfvwvl-NH_2	803
VRC01-ReH4	Ac-vfvwvy-NH_2	853
VRC01-ReH5	Ac-vfvwel-NH_2	833
VRC01-ReH6	Ac-efvyvl-NH_2	810
VRC01-ReH7	Ac-efvyvy-NH_2	860
VRC01-ReH8	Ac-efvyel-NH_2	840
VRC01-ReH9	Ac-efvmvl-NH_2	778
VRC01-ReH10	Ac-efvmvy-NH_2	828
VRC01-ReH11	Ac-efvwvl-NH_2	833
VRC01-ReH12	Ac-efvwvy-NH_2	883
VRC01-ReH13	Ac-efvwel-NH_2	863
VRC01-ReH14	Ac-efvwey-NH_2	913
VRC01-ReH11-Bio	Bio-AOA-efvwel-NH_2	1192
VRC01-ReH12-Bio	Bio-AOA-efvwvy-NH_2	1212
VRC01-ReH13-Bio	Bio-AOA-efvwvl-NH_2	1162
VRC01-ReH15	Ac-EFVWEL-NH_2	863
VRC01-ReH16	Ac-EFVWVY-NH_2	883
VRC01-ReH17	Ac-EFVWVL-NH_2	833
V3-$Loop_{HXBc2}$	Biotin-XGX[CTRPNNNTRKRIRIQRGPGRAFVTIGKIGNMRQAHC]-NH_2	4633
Fluo-V3-$Loop_{HXBc2}$*	5(6)-carboxyfluorescein-AOA-[CTRPNNNTRKRIRIQRGPGRAFVTIGKIGNMRQAHC]-NH_2	4826
Disulfide $Loop_{HXBc2}$-ox	Bio-XKDQQLLGIWG[CSGKLIC]TTAV-NH_2	2571

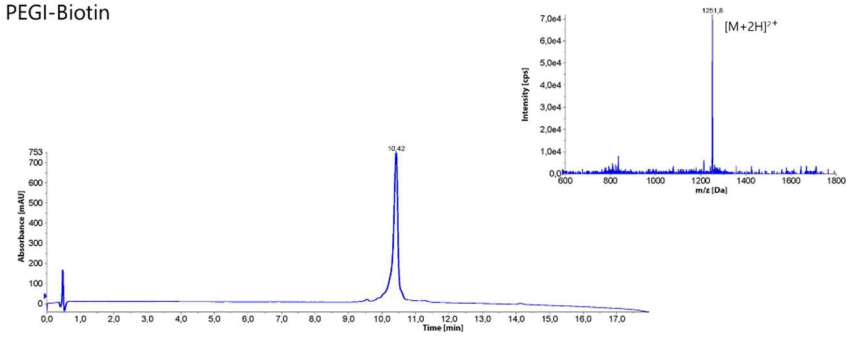
* synthesis was performed by Karen Fiebig and peak purity >90 % was confirmed by LC-MS

LC-MS data

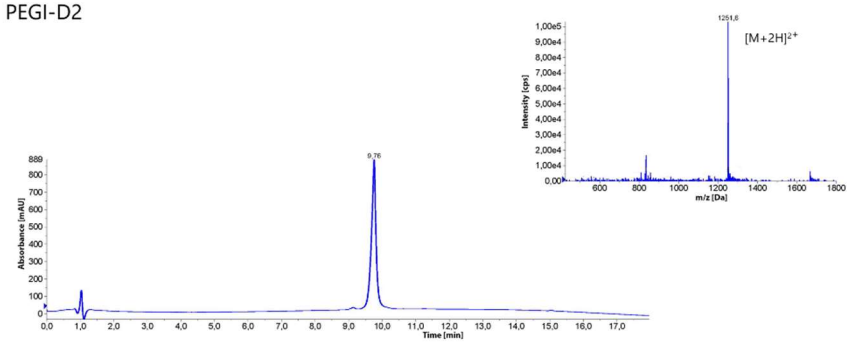
PEGI



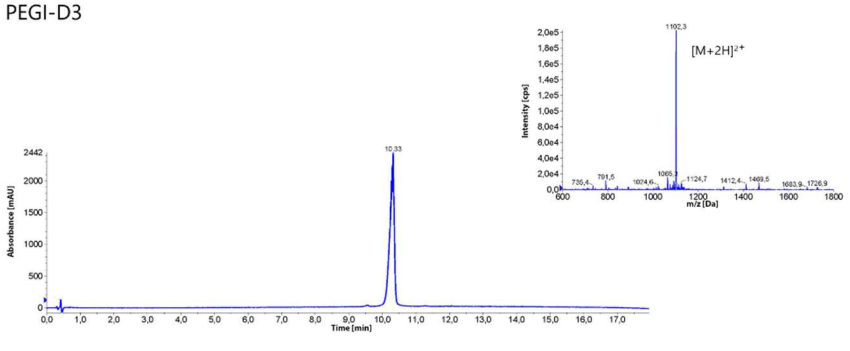
PEGI-Biotin



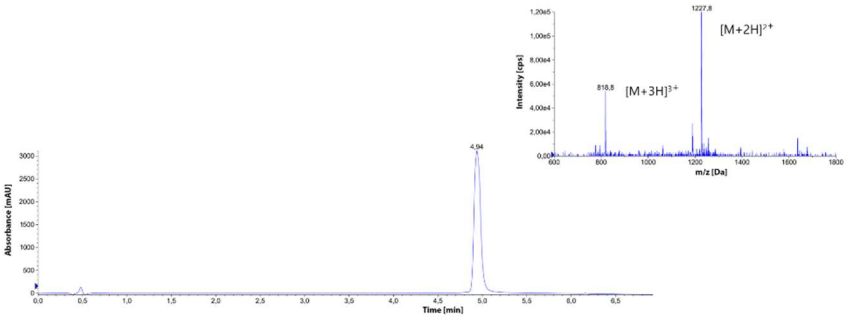
PEGI-D2



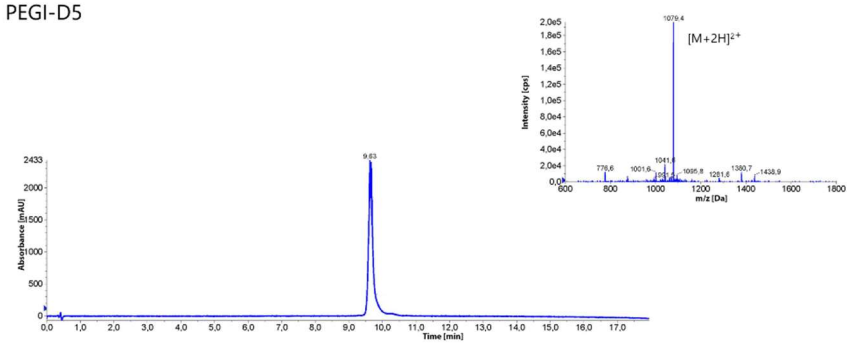
PEGI-D3



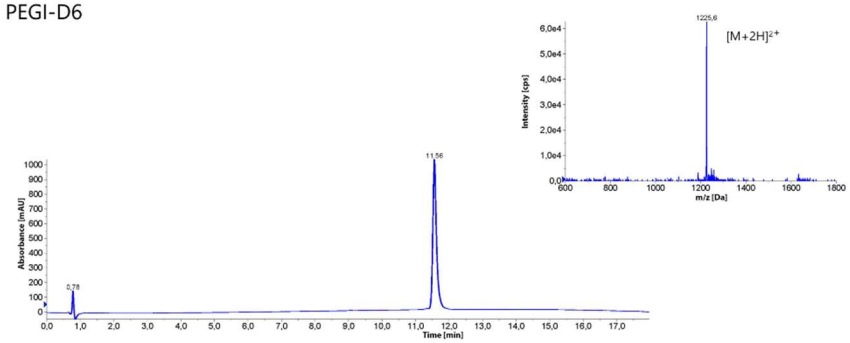
PEGI-D4



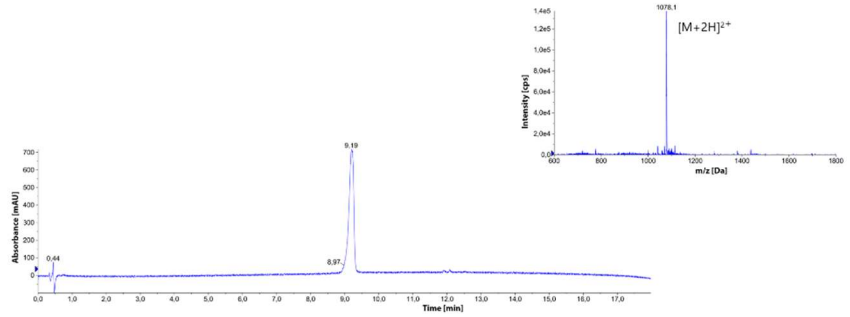
PEGI-D5



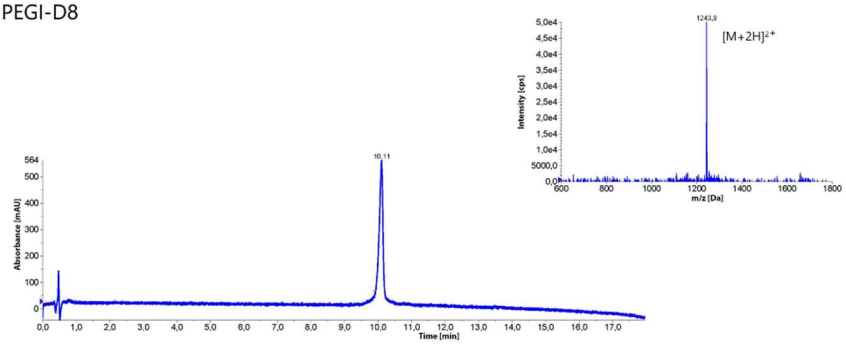
PEGI-D6



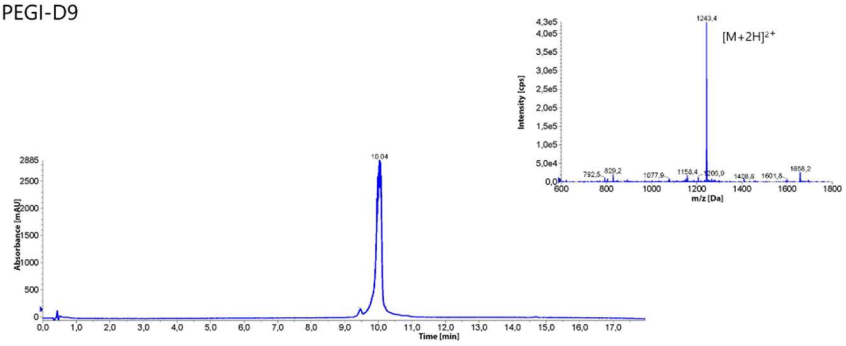
PEGI-D7



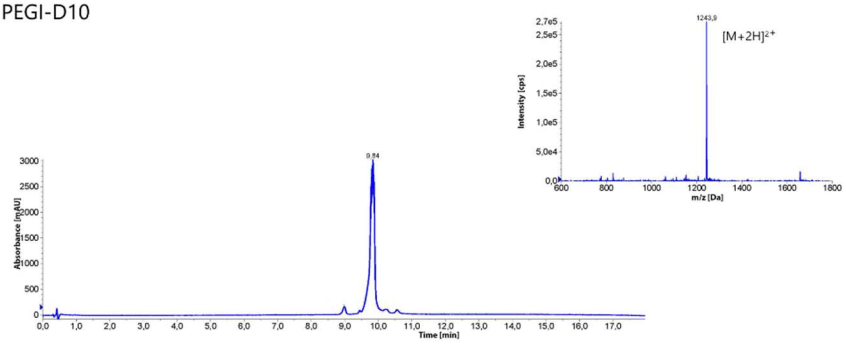
PEGI-D8



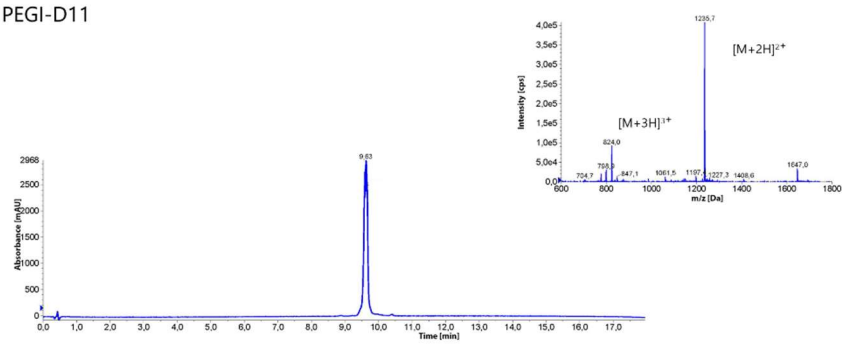
PEGI-D9



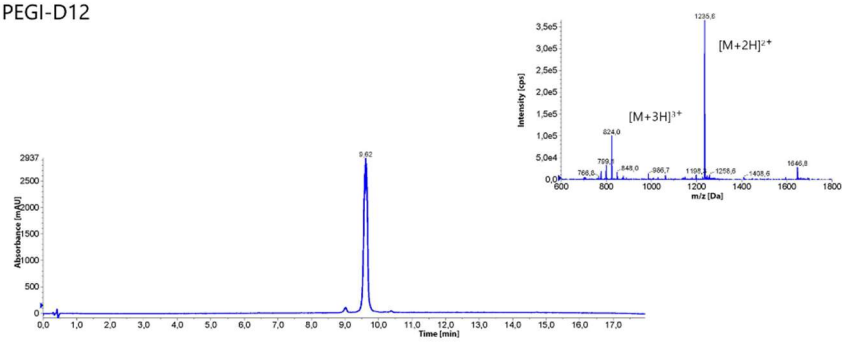
PEGI-D10



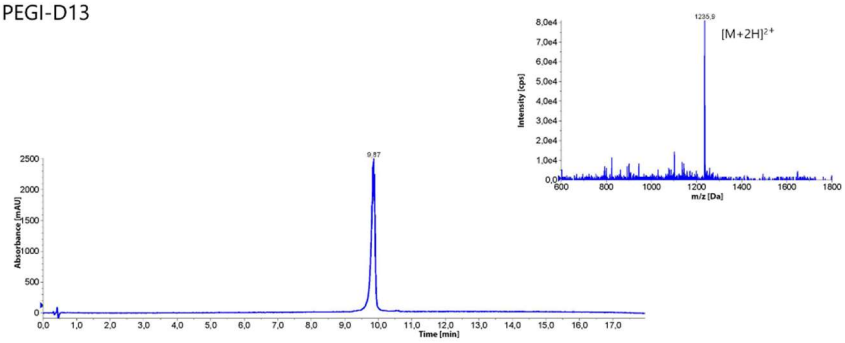
PEGI-D11



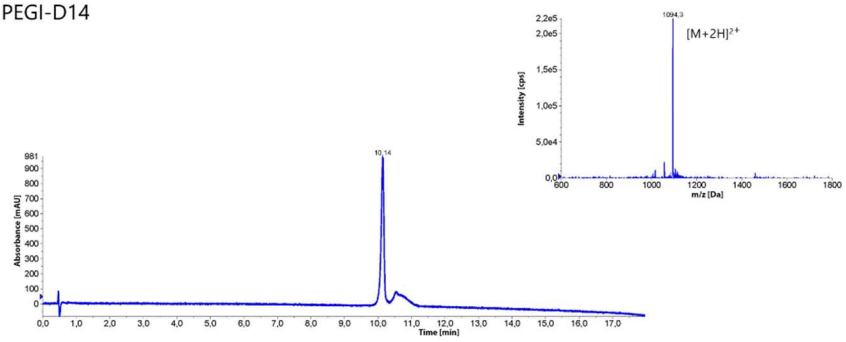
PEGI-D12



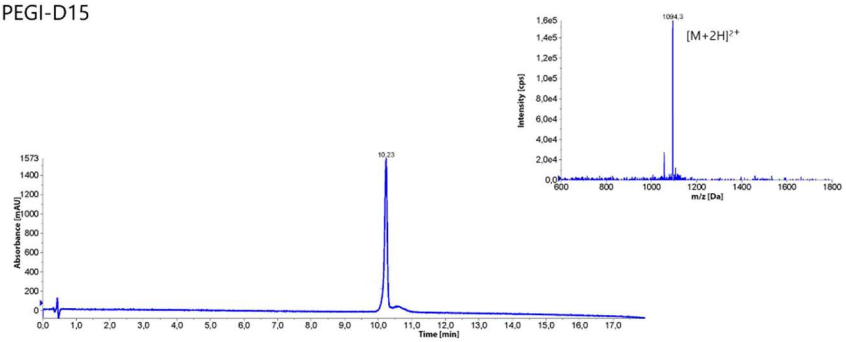
PEGI-D13



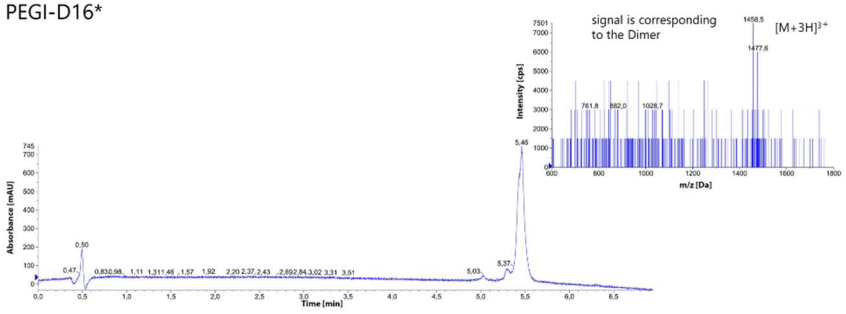
PEGI-D14



PEGI-D15

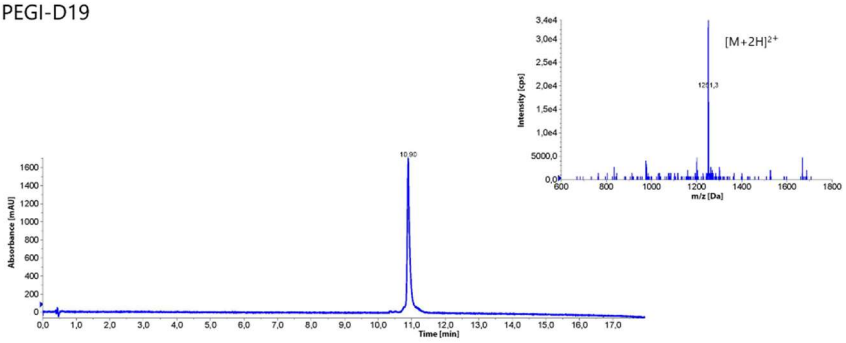


PEGI-D16*

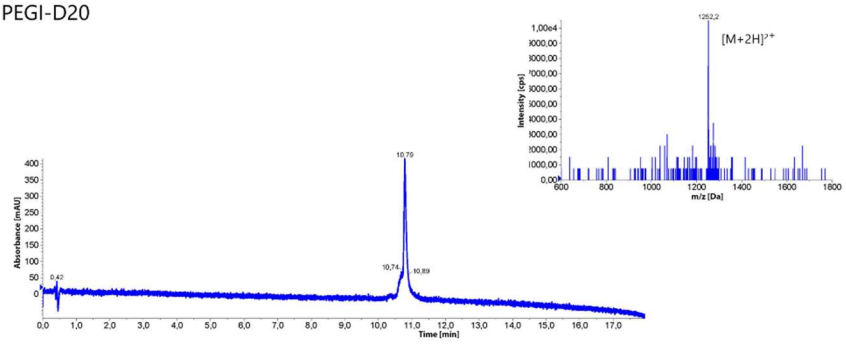


* LC-MS was measured with the LC-MS method purification. During Purification dimerization occurred and the oxidized monomer was not obtained.

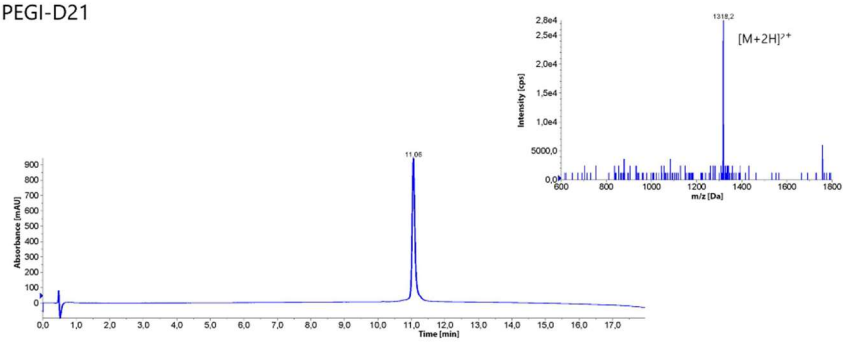
PEGI-D19



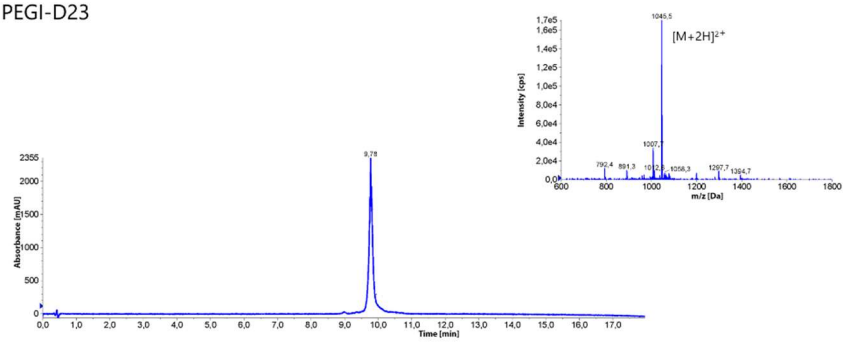
PEGI-D20



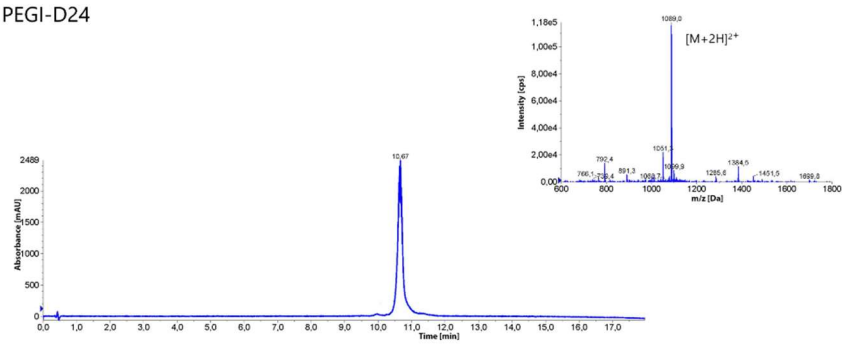
PEGI-D21



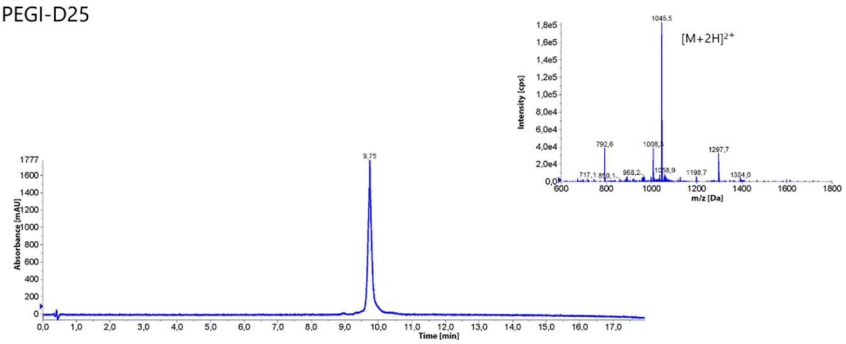
PEGI-D23



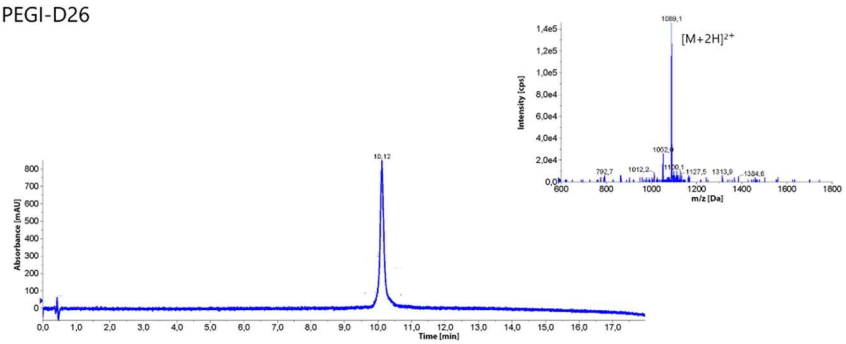
PEGI-D24



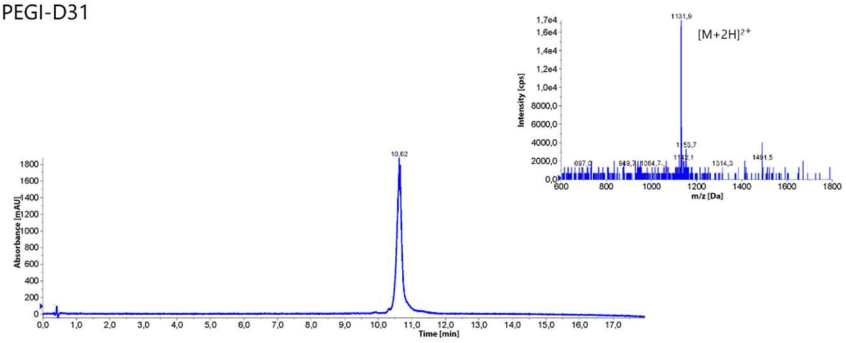
PEGI-D25



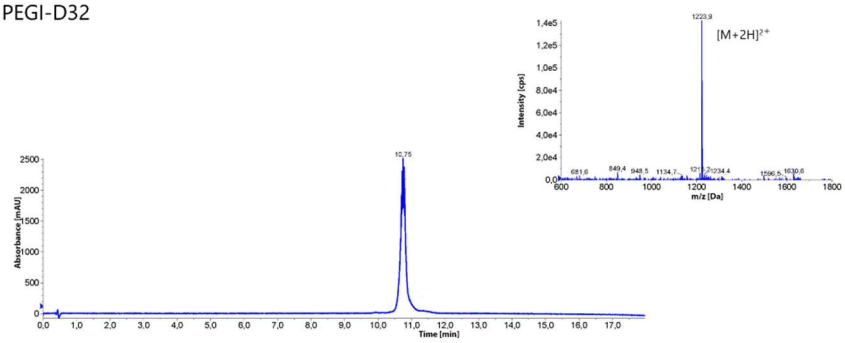
PEGI-D26



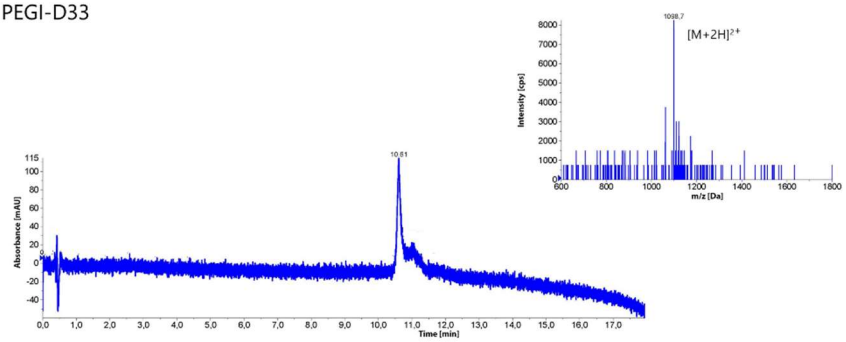
PEGI-D31



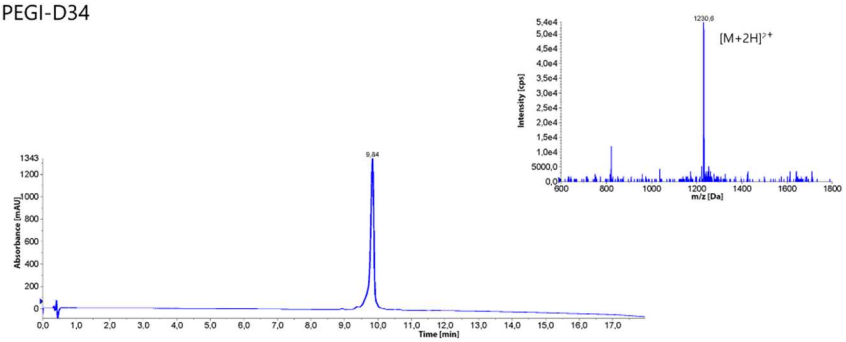
PEGI-D32



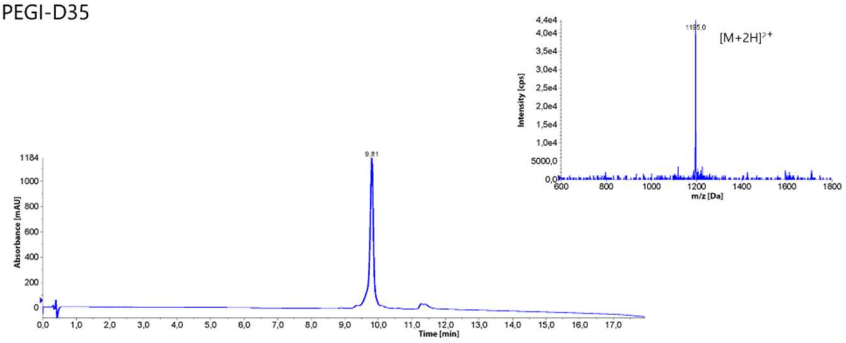
PEGI-D33



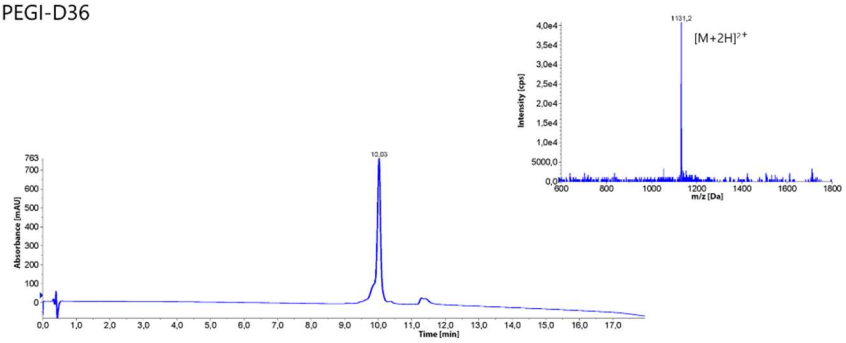
PEGI-D34



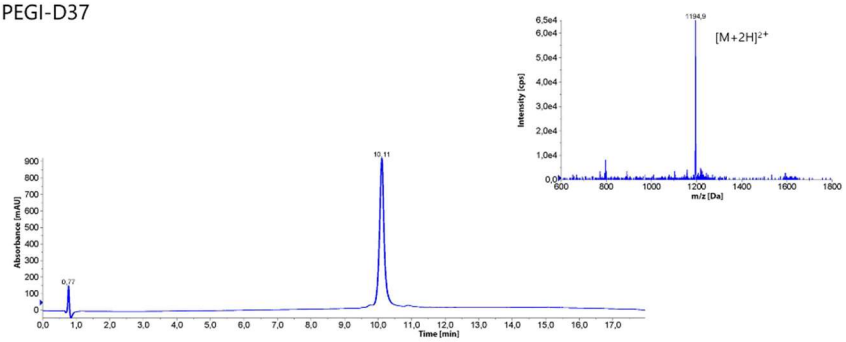
PEGI-D35



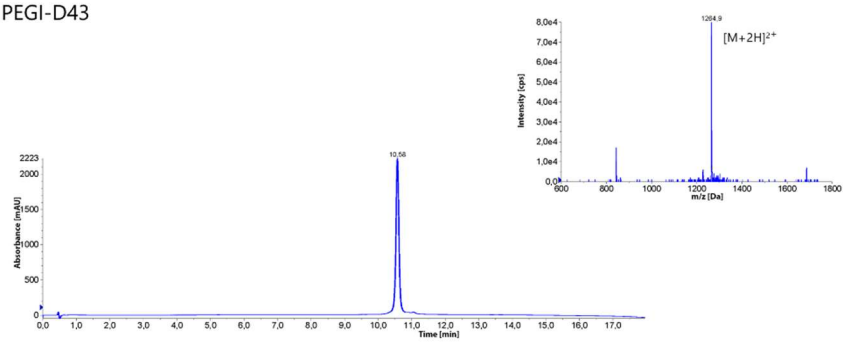
PEGI-D36



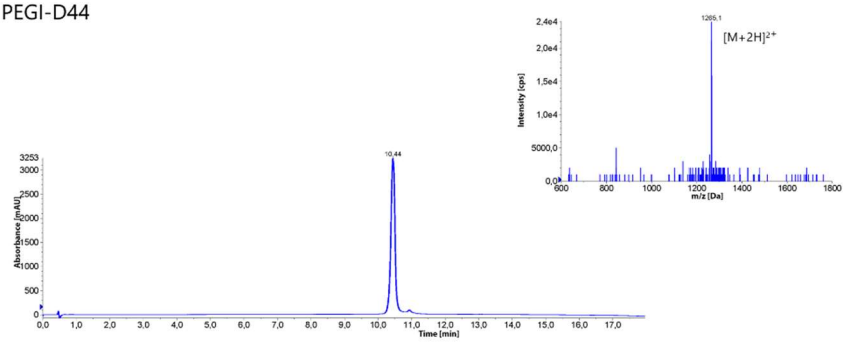
PEGI-D37



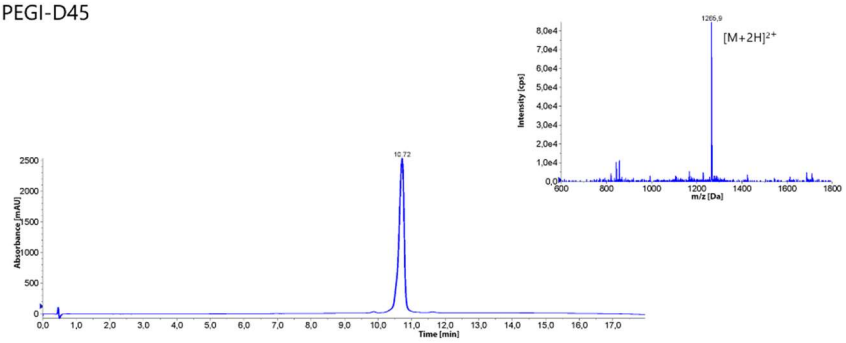
PEGI-D43



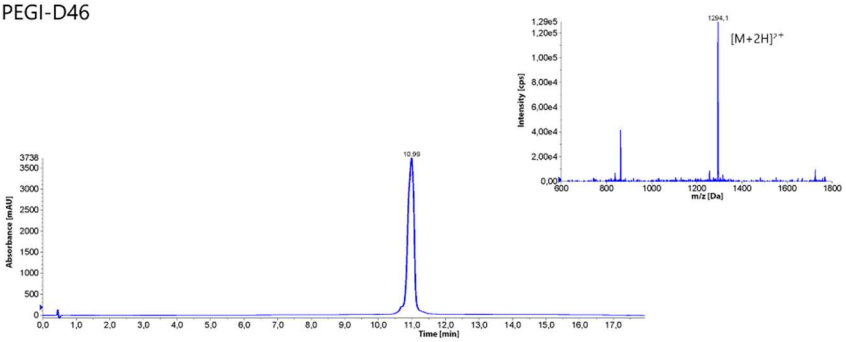
PEGI-D44



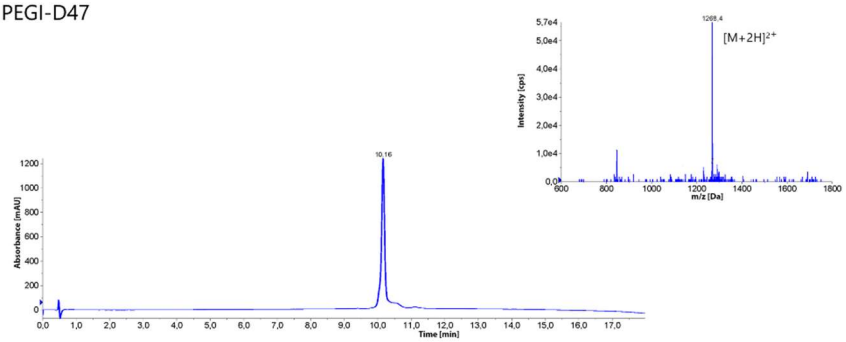
PEGI-D45



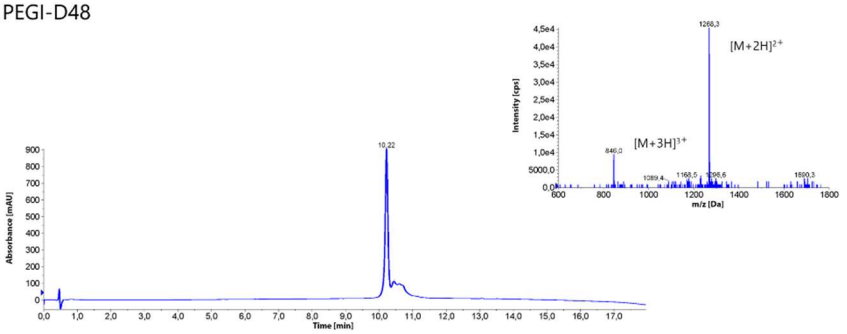
PEGI-D46



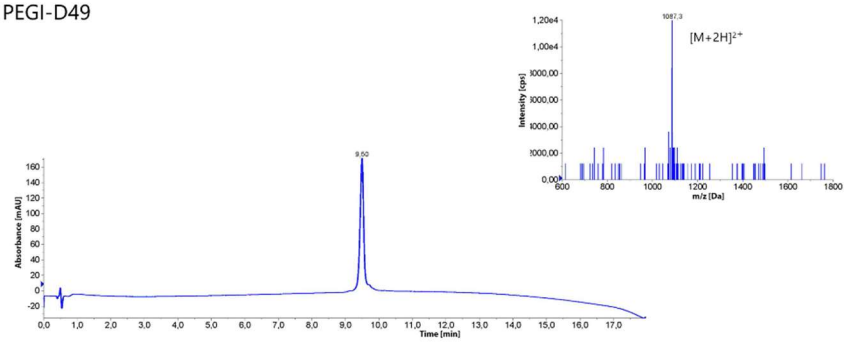
PEGI-D47



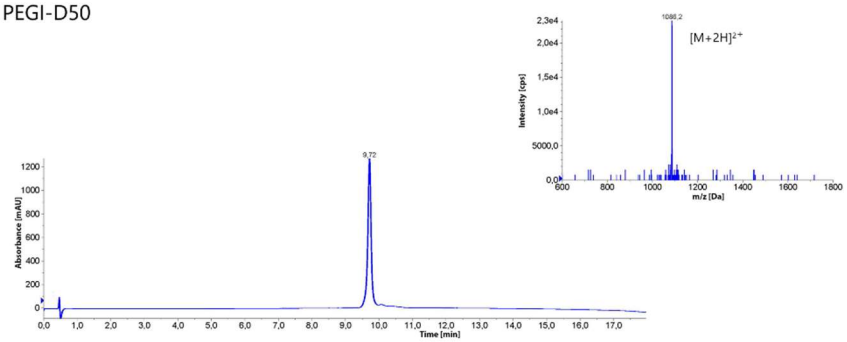
PEGI-D48



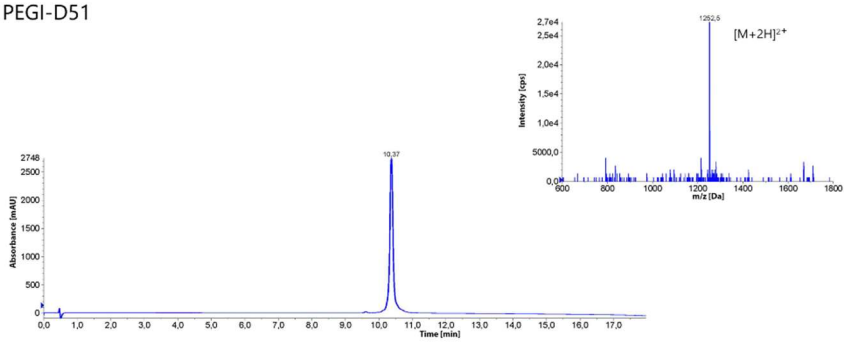
PEGI-D49



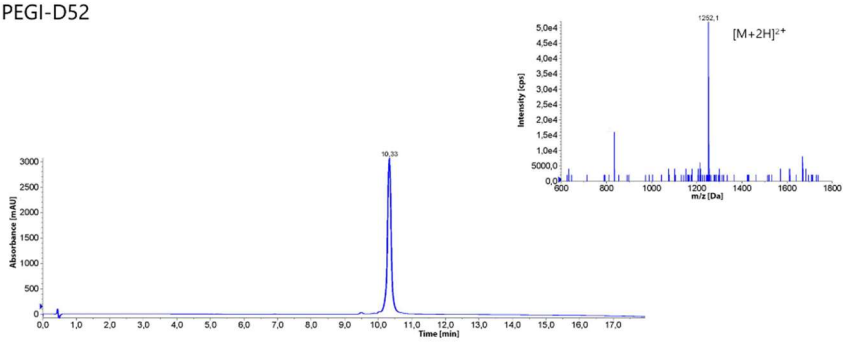
PEGI-D50



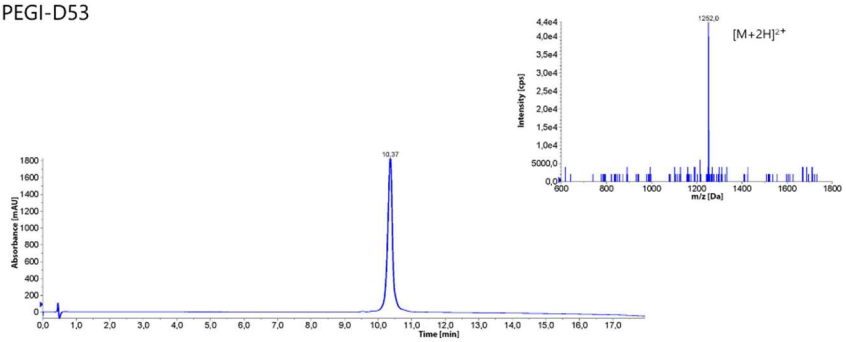
PEGI-D51



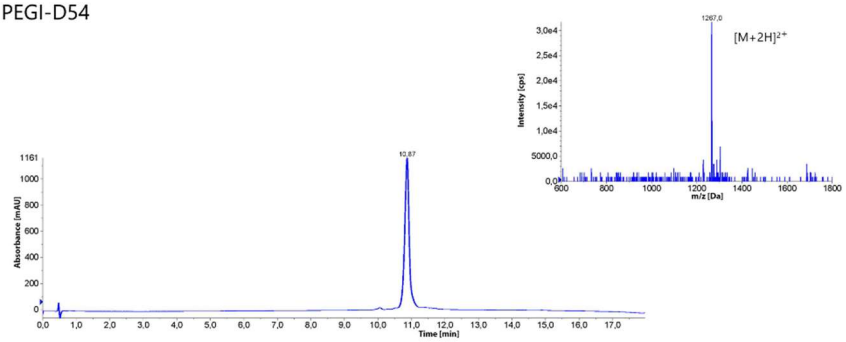
PEGI-D52



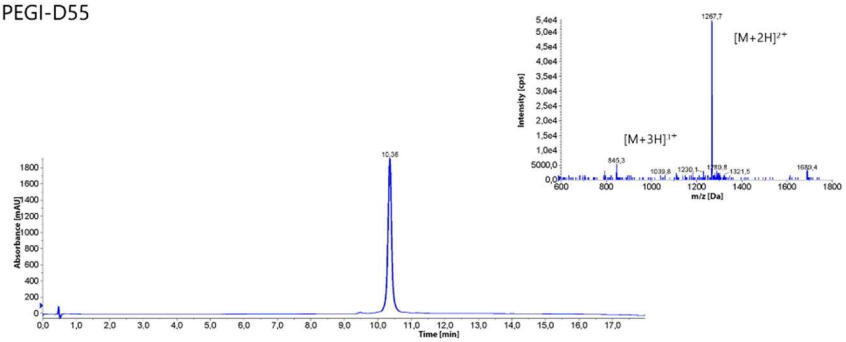
PEGI-D53



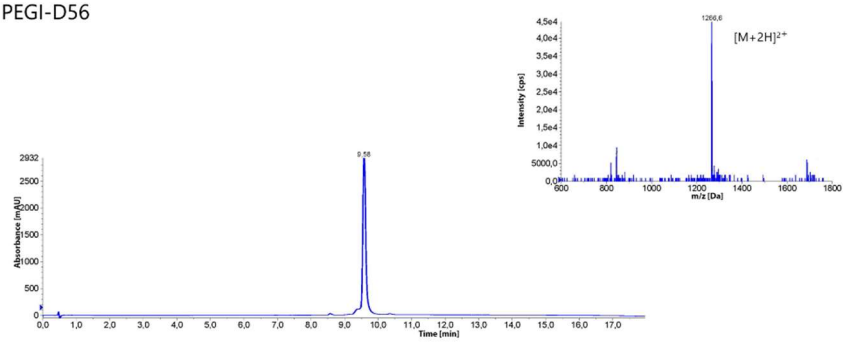
PEGI-D54



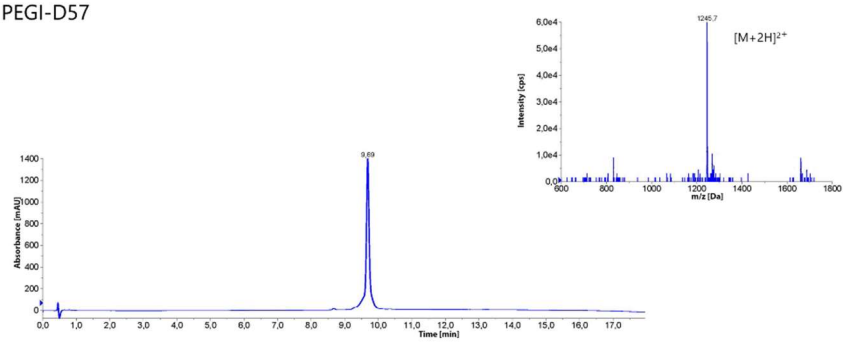
PEGI-D55



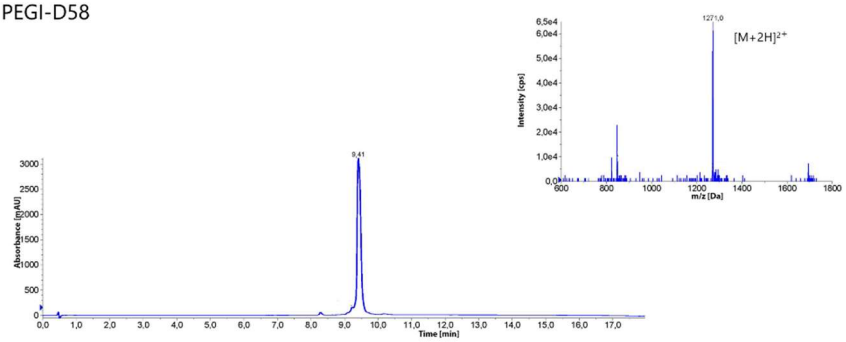
PEGI-D56



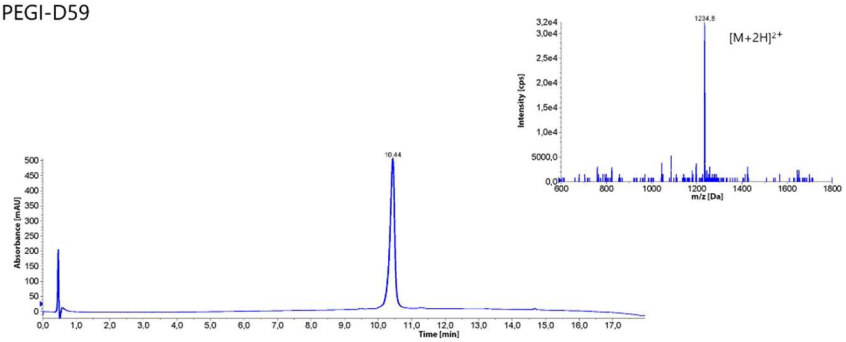
PEGI-D57



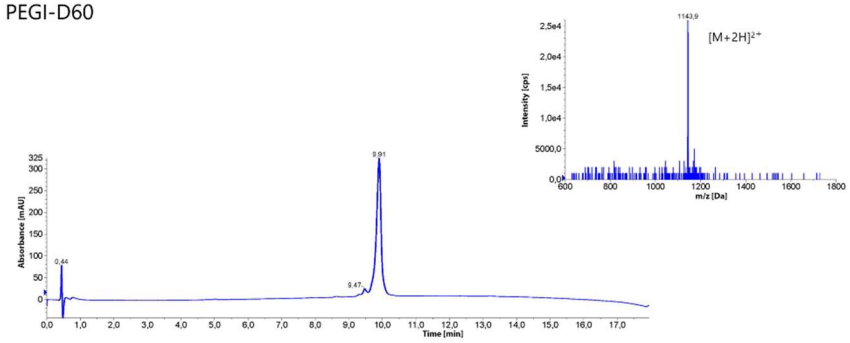
PEGI-D58



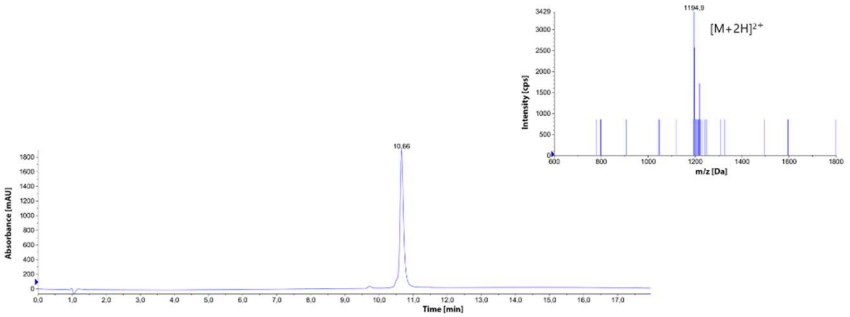
PEGI-D59



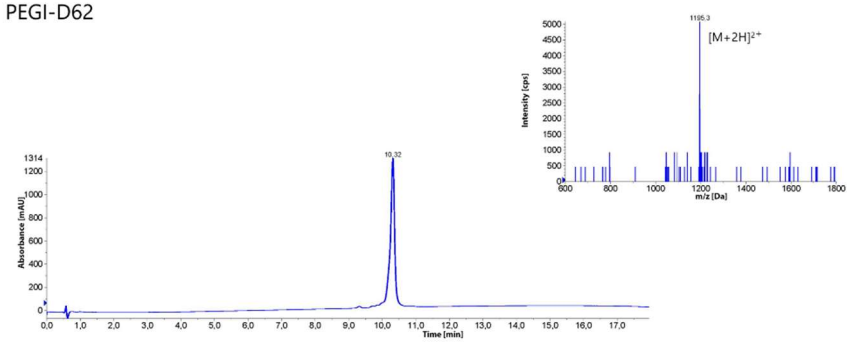
PEGI-D60



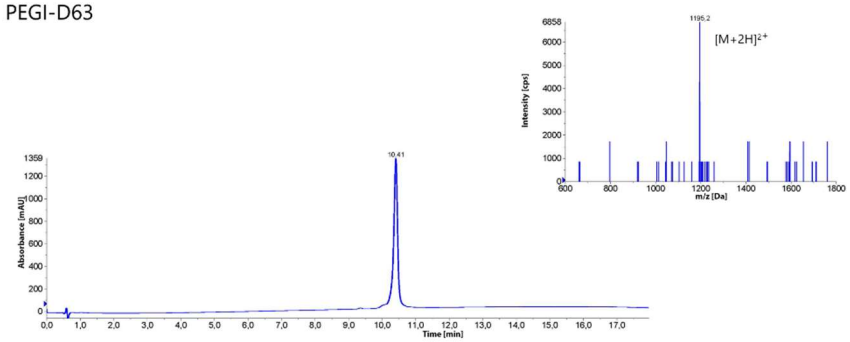
PEGI-D61



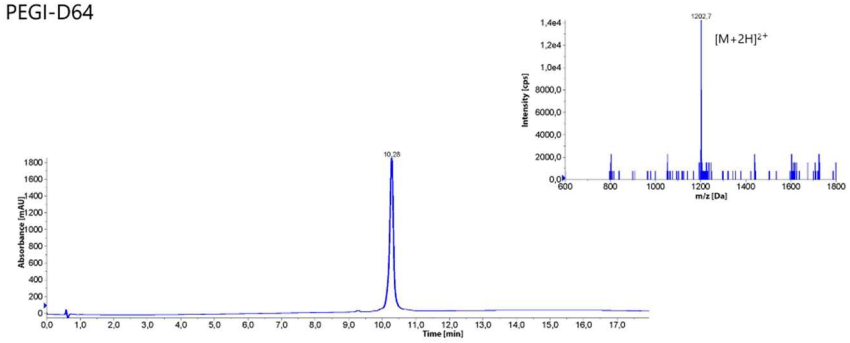
PEGI-D62



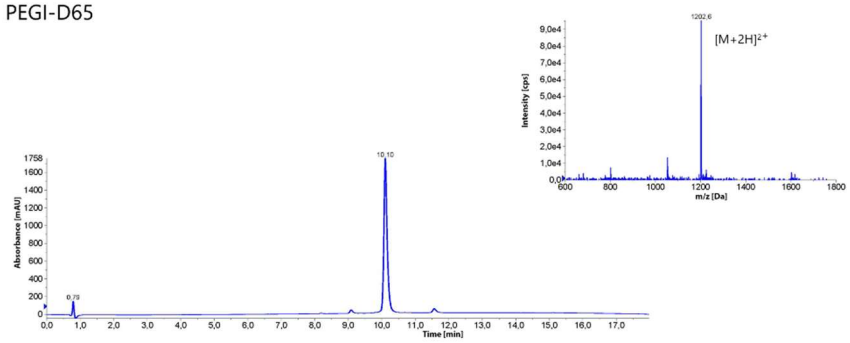
PEGI-D63



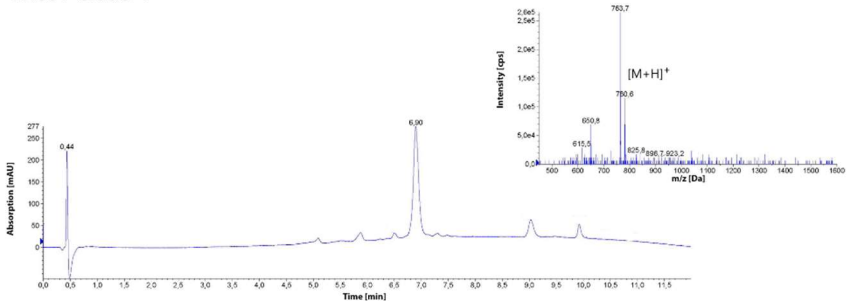
PEGI-D64



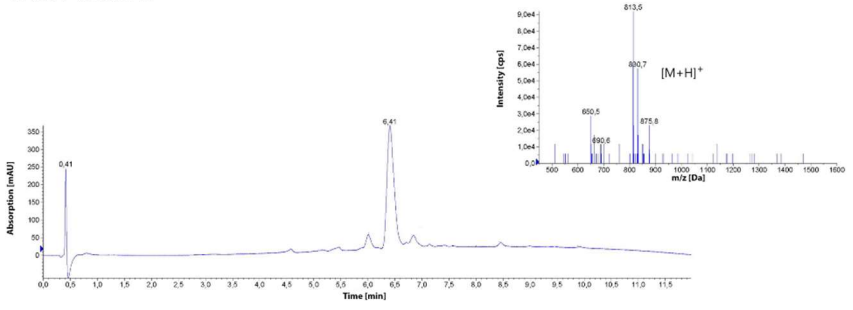
PEGI-D65



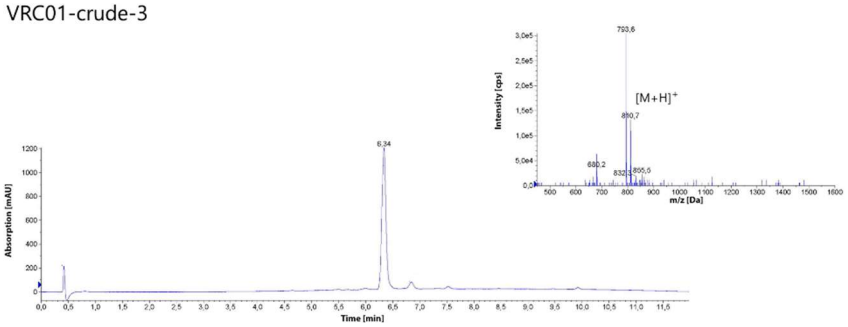
VRC01-crude-1



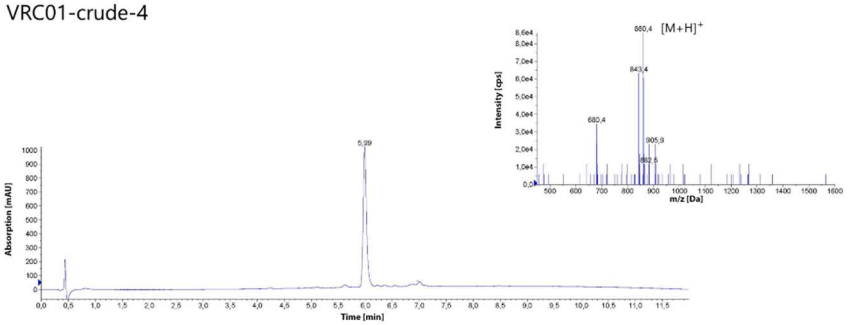
VRC01-crude-2



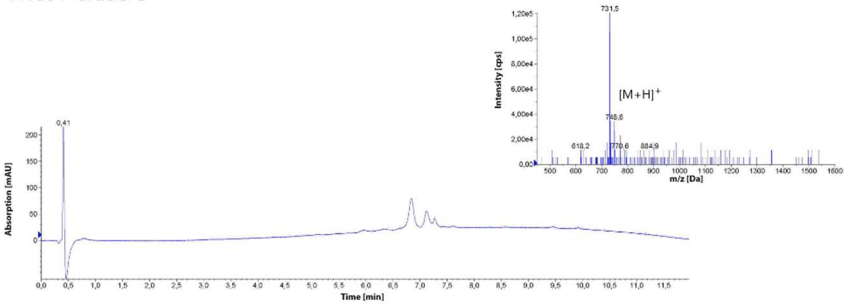
VRC01-crude-3



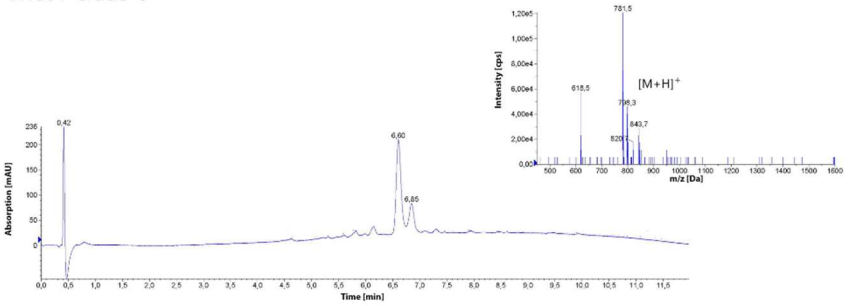
VRC01-crude-4



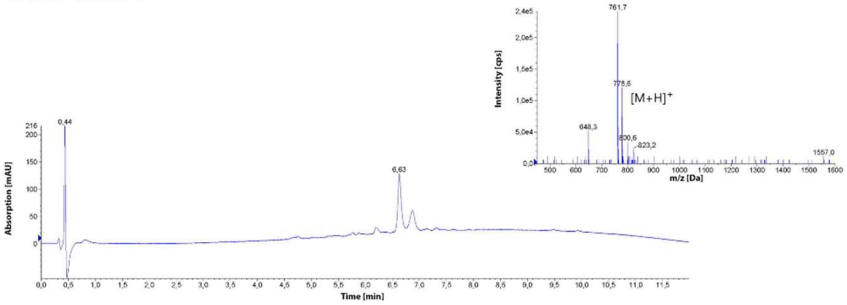
VRC01-crude-5



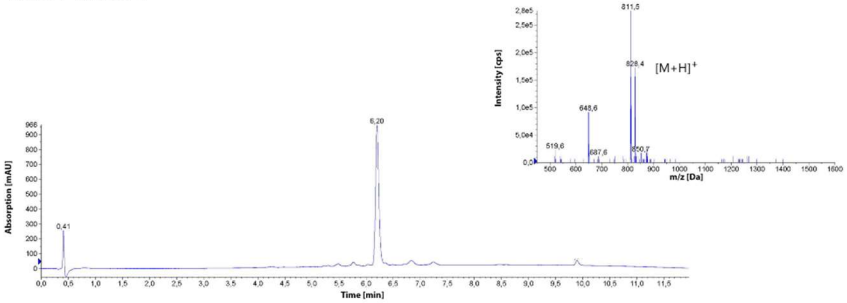
VRC01-crude-6



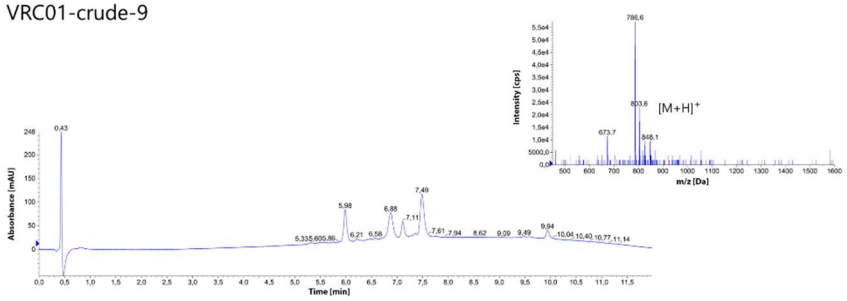
VRC01-crude-7



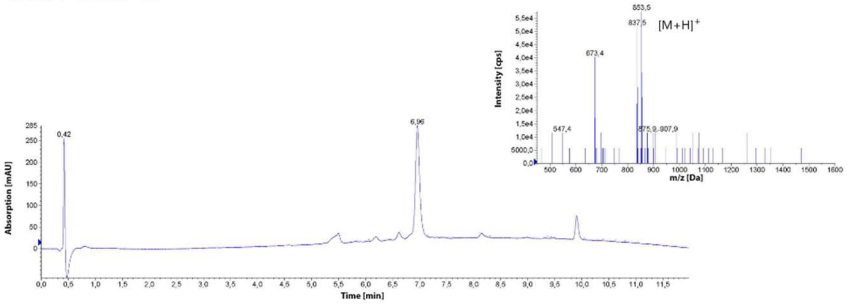
VRC01-crude-8



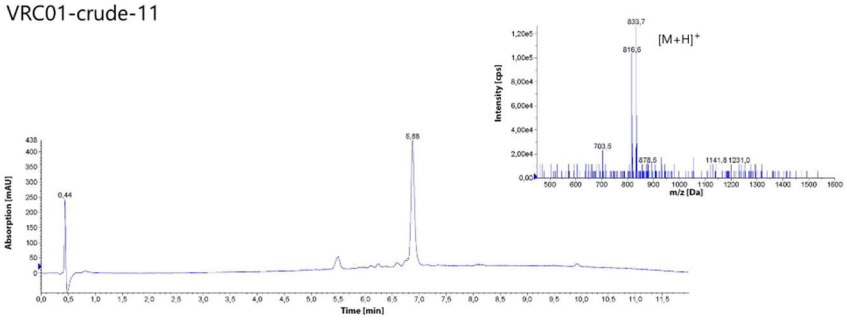
VRC01-crude-9



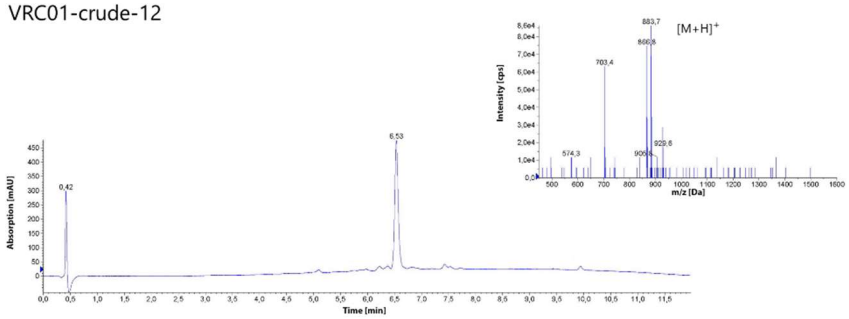
VRC01-crude-10



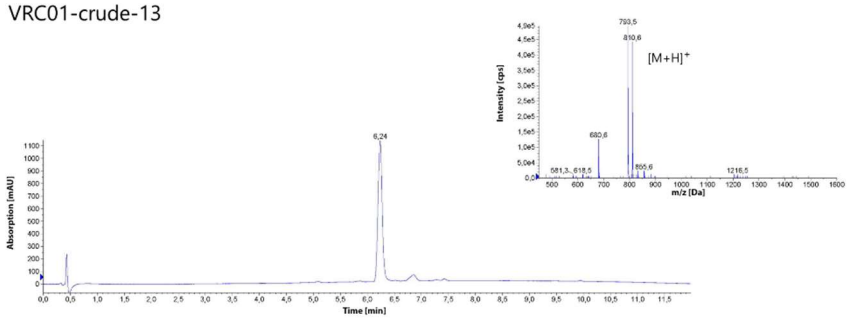
VRC01-crude-11



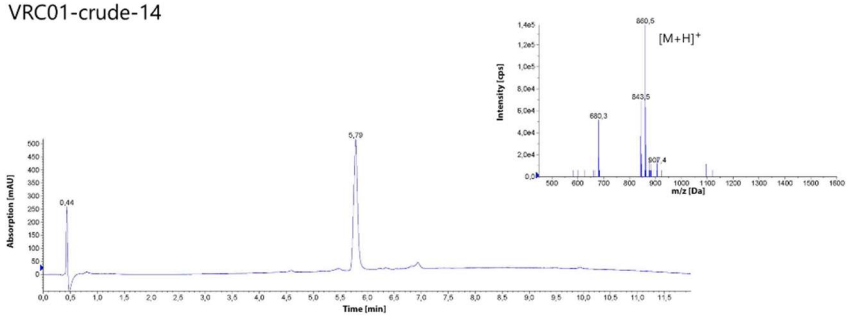
VRC01-crude-12



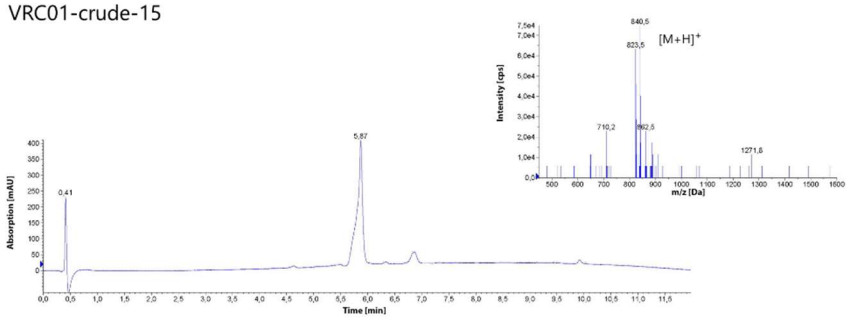
VRC01-crude-13



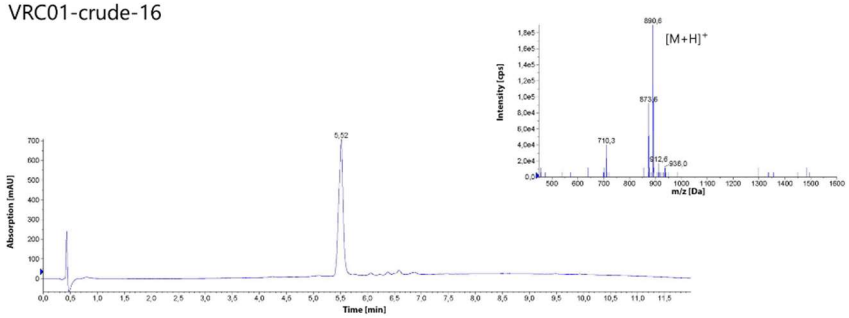
VRC01-crude-14



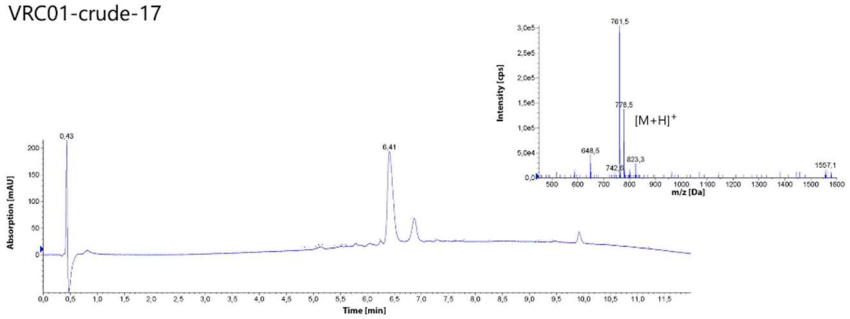
VRC01-crude-15



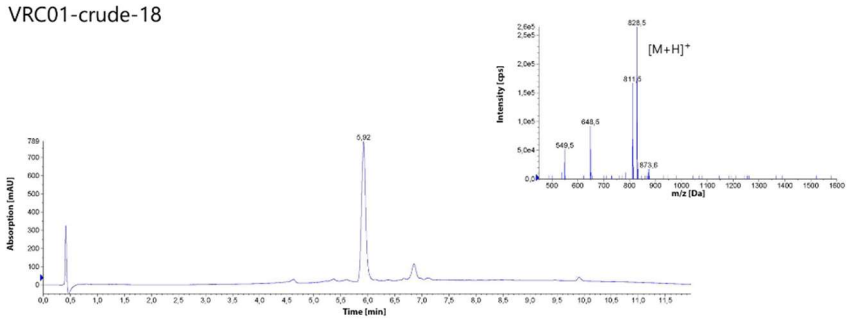
VRC01-crude-16



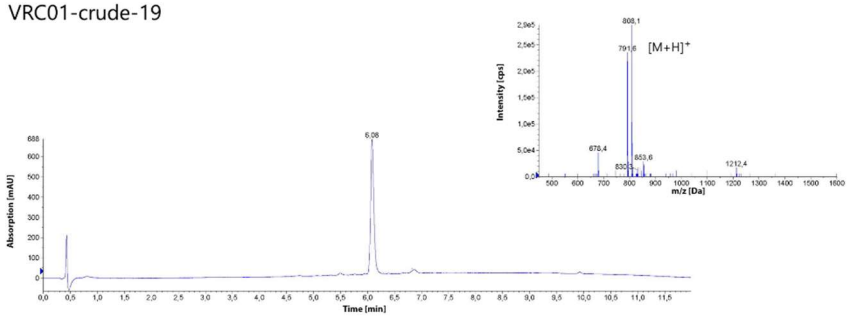
VRC01-crude-17



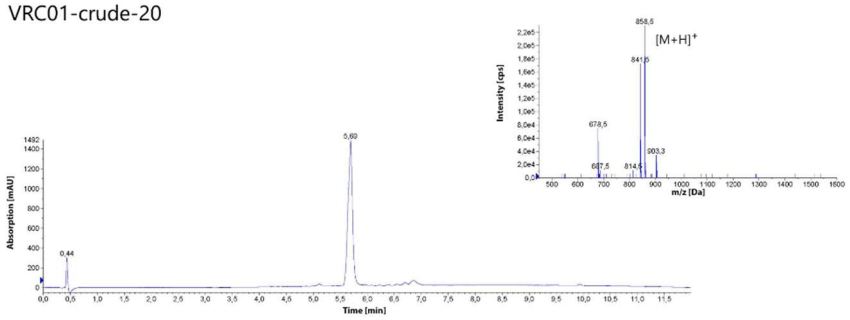
VRC01-crude-18



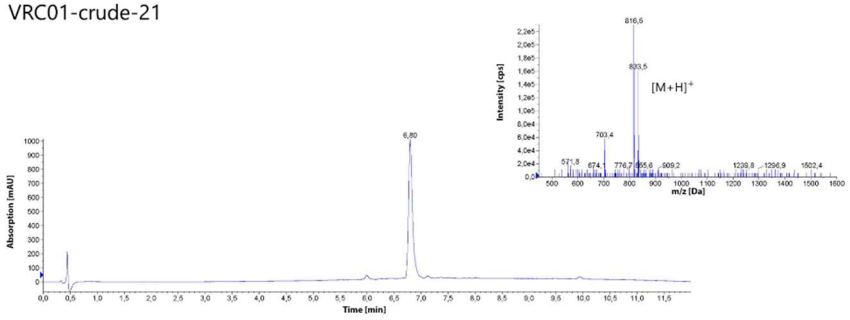
VRC01-crude-19



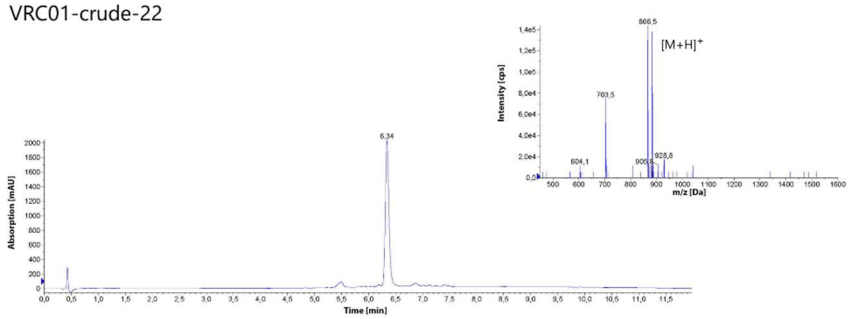
VRC01-crude-20



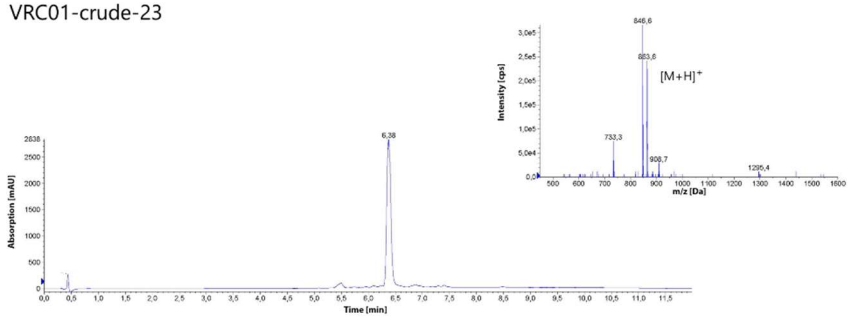
VRC01-crude-21



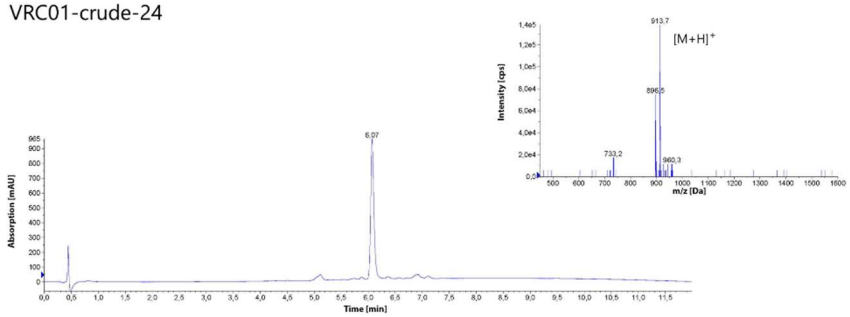
VRC01-crude-22



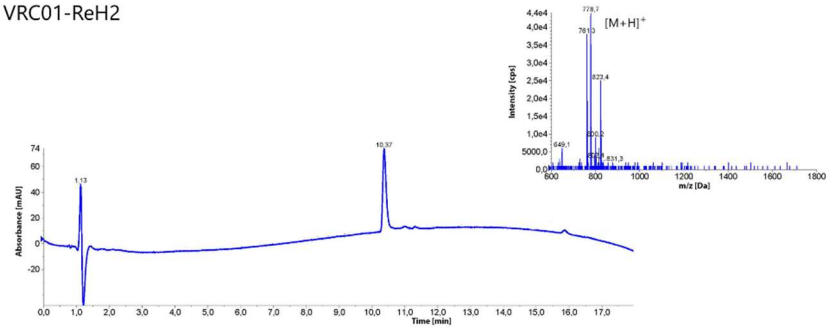
VRC01-crude-23



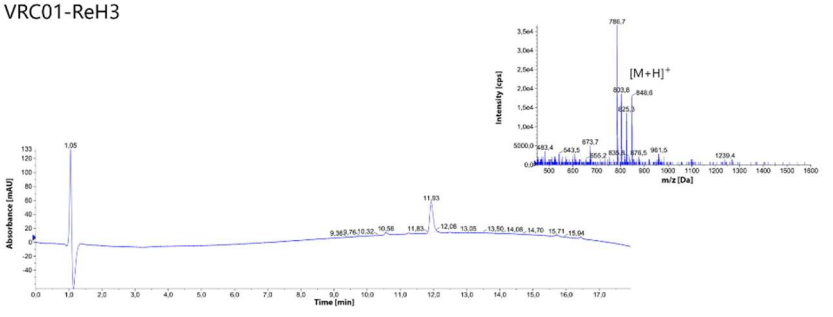
VRC01-crude-24



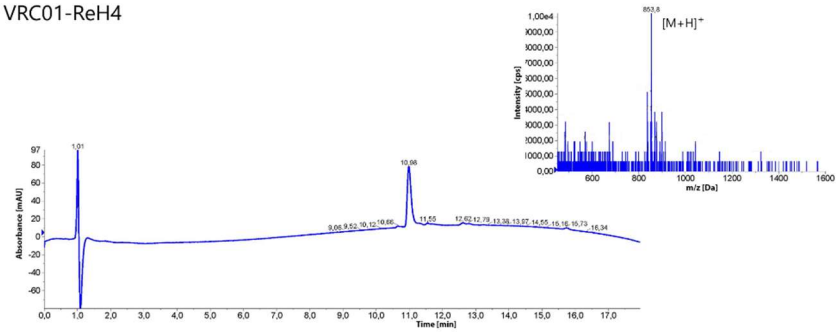
VRC01-ReH2



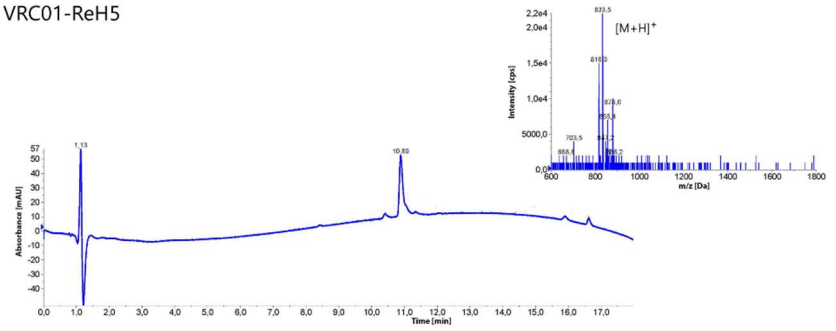
VRC01-ReH3



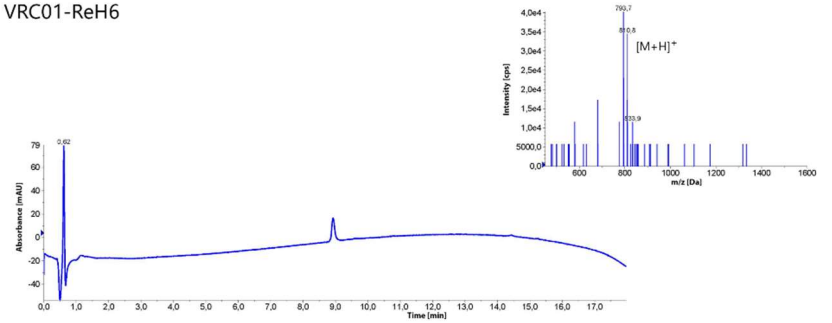
VRC01-ReH4



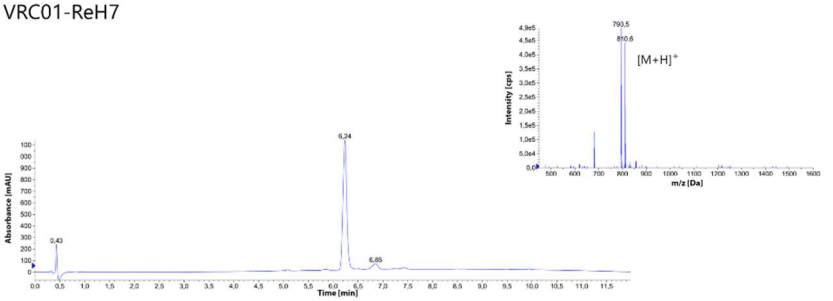
VRC01-ReH5



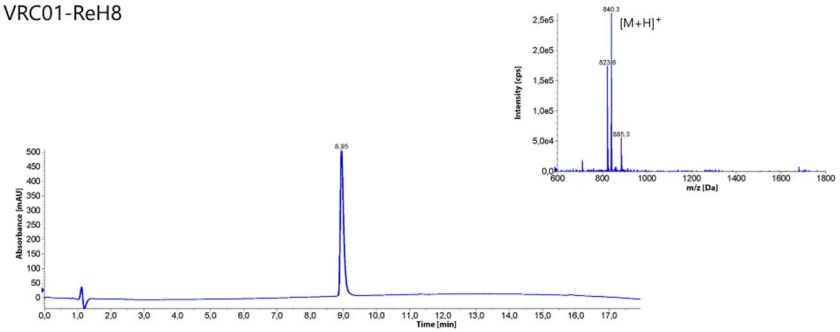
VRC01-ReH6



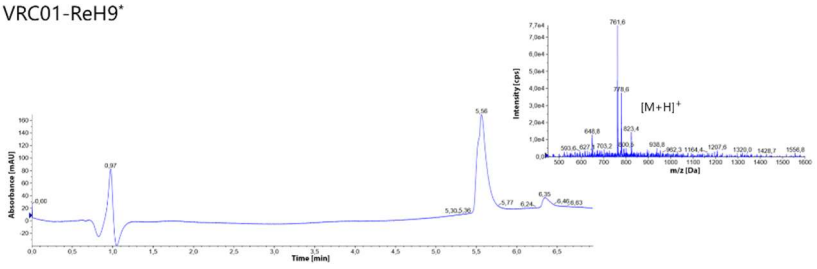
VRC01-ReH7



VRC01-ReH8

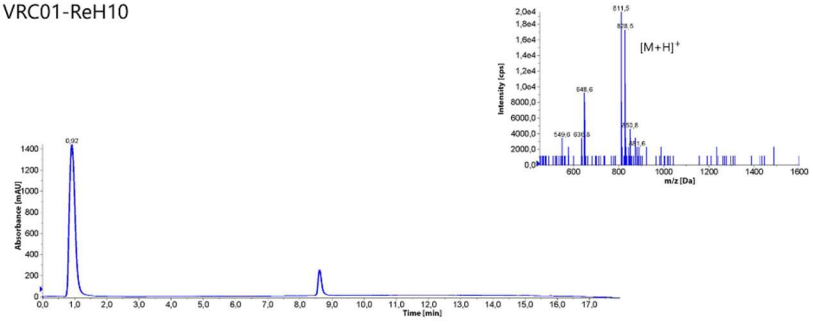


VRC01-ReH9*

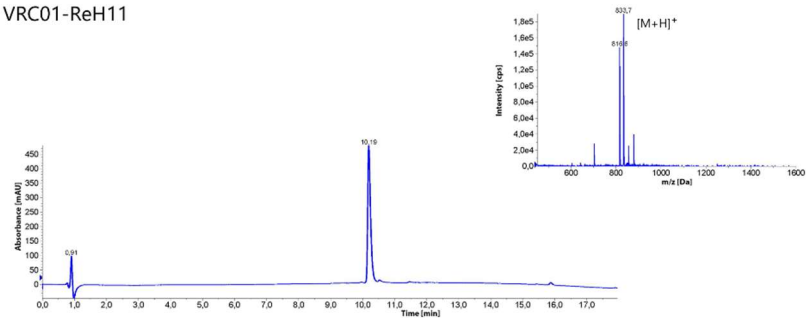


* LC-MS was measured with the LC-MS method purification

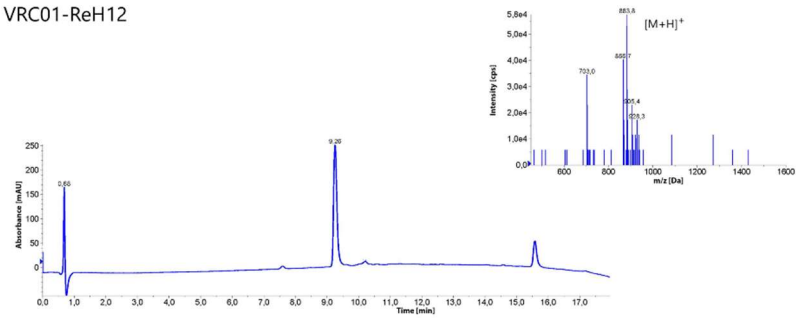
VRC01-ReH10



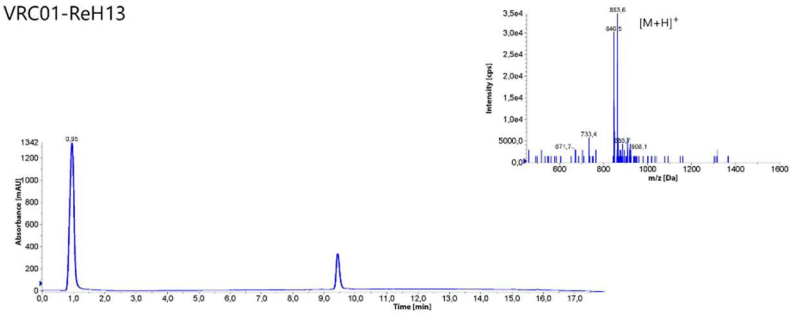
VRC01-ReH11



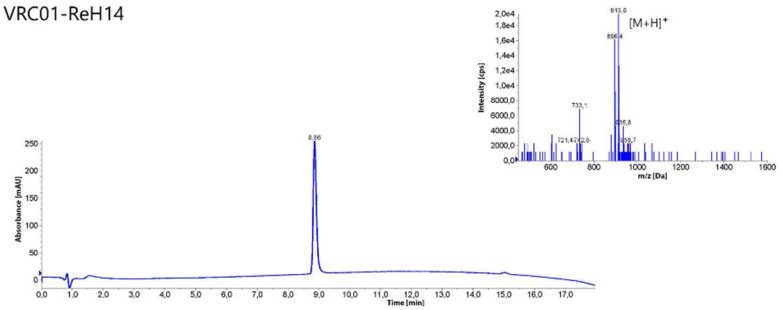
VRC01-ReH12



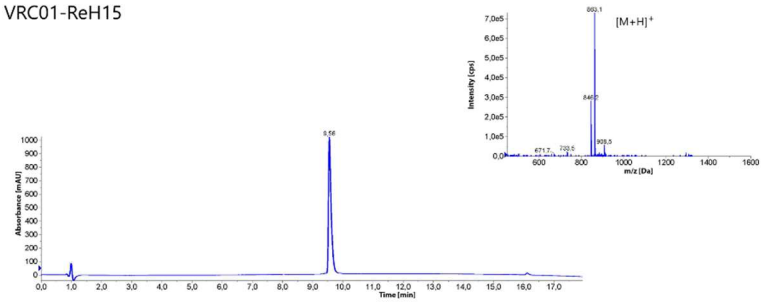
VRC01-ReH13



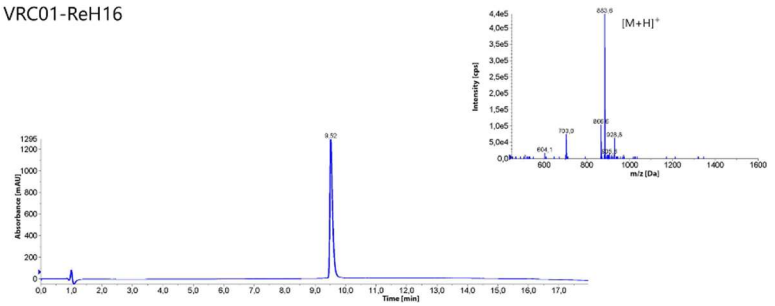
VRC01-ReH14



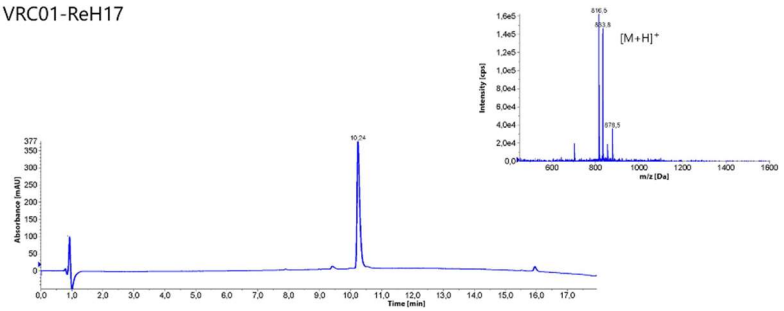
VRC01-ReH15



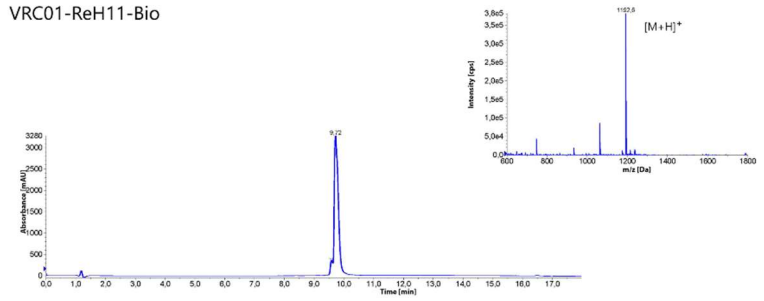
VRC01-ReH16



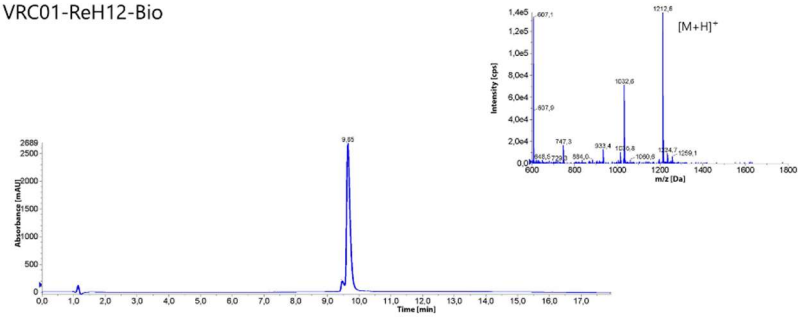
VRC01-ReH17



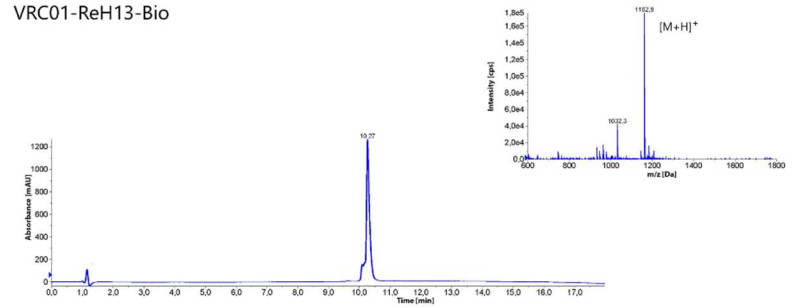
VRC01-ReH11-Bio



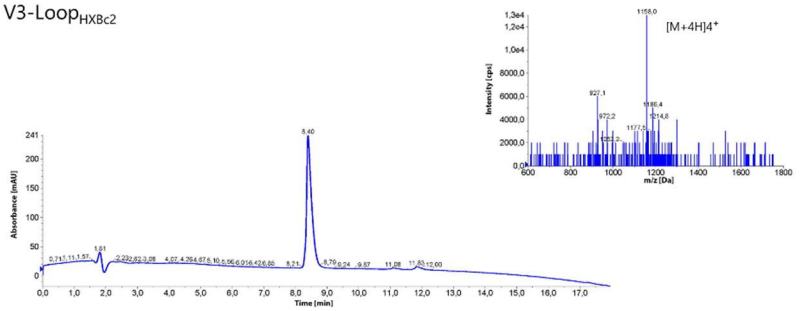
VRC01-ReH12-Bio



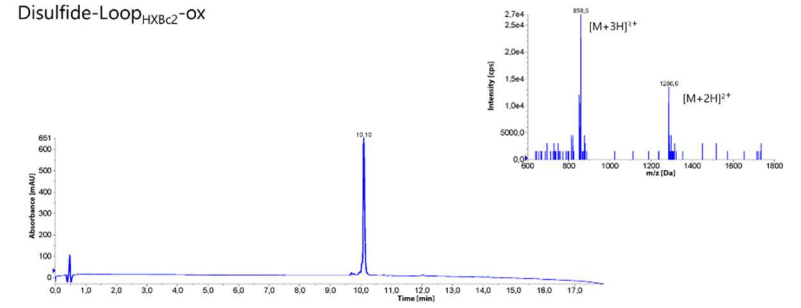
VRC01-ReH13-Bio



V3-Loop$_{HXBc2}$



Disulfide-Loop$_{HXBc2}$-ox



7.2 Supplementary data

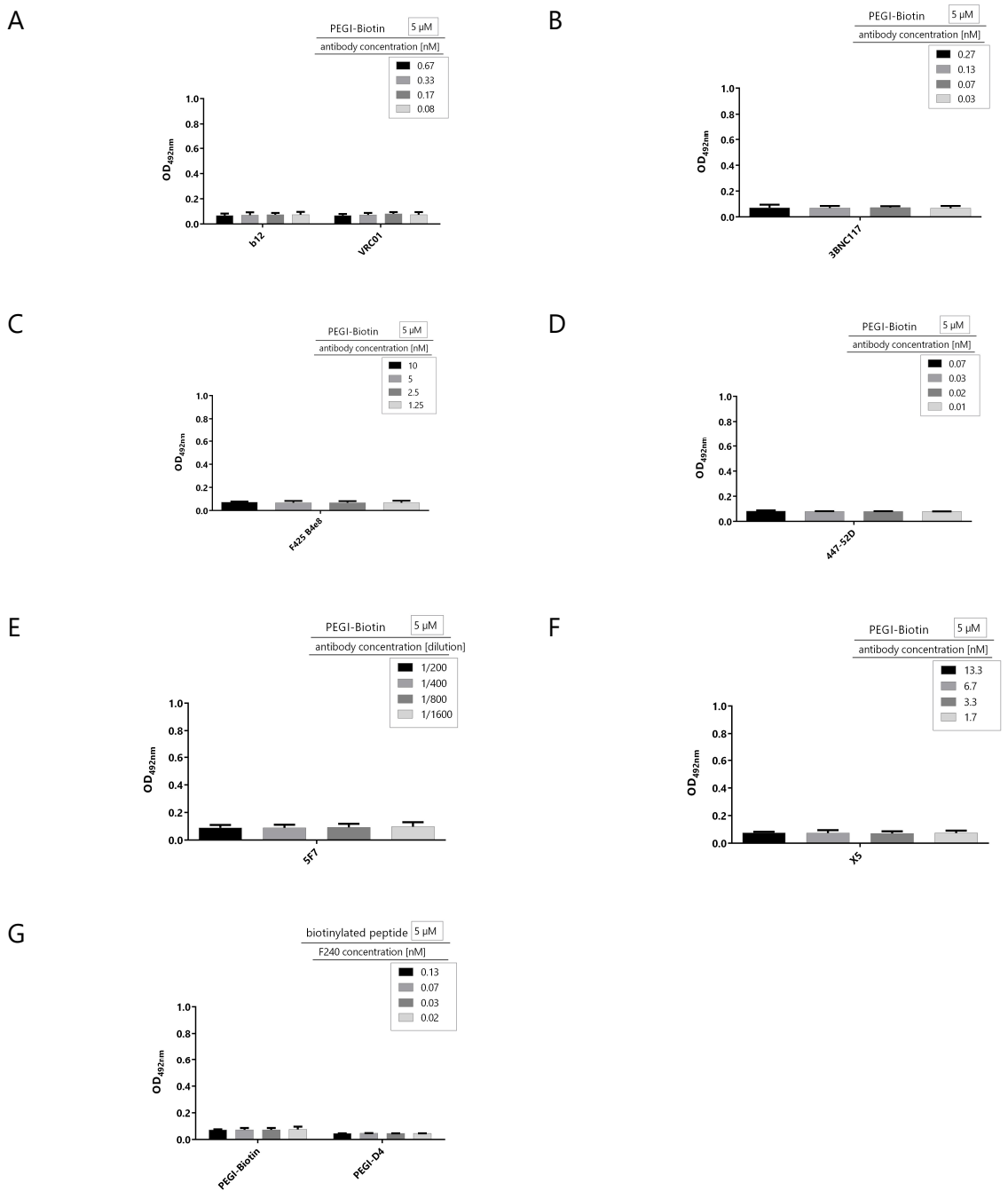


Figure 65: Binding of PEGI-Biotin and PEGI-D4 to different HIV-1 antibody, A: Binding to b12 and VRC01. B: Binding to 3BNC117. C: Binding to F425 B4e8. D: Binding to 447-52D. E: Binding to 5F7. F: Binding to X5: G: Binding to F240, tested in Binding Assay Format B. Four concentrations were tested per antibody and presented in bar graphs.

Table 42: Overview of HIV-1 strain panel and results of HIV-1 inhibition by T20, tested in Luciferase assay. n=1.

group	IC_{50} range [nM]
I	0.5 - 10
II	10 - 100
III	100 - 500

HIV-1 strains				T20
Abbreviation	Isolate	HIV-1-subtype	Tropism	IC_{50} group
A1	92UG029	A	X4	I
A2	00KE_KER2018	A	dual	III
AG01	CM.0005BBY	CRF02_AG	R5	II
AG02	01CM.0008BBY	CRF02_AG	R5	II
B2	90TH_BK132	B	X4	II
SF33	SF33	B	X4	I
MN	MN	B	X4	I
HxBC2	HxB2	B	X4	I
NL4-3	NL4-3	B	X4	III
BAL	93BR093	B	R5	I
C1	92_BR025	C	R5	II
C2	99_ET14	C	R5	II
D1	92UG021	D	X4	III
D3	92UG024	D	X4	I
F1	93BR029	F	R5	I
F2	93BR020	F	dual	I
G ARP	RU570	G	R5	I
G/H	VI525	G/H	R5	I
O BCF06	O BCF06	group O	X4	I
O BCF011	O BCF011	group O	X4	n.d.

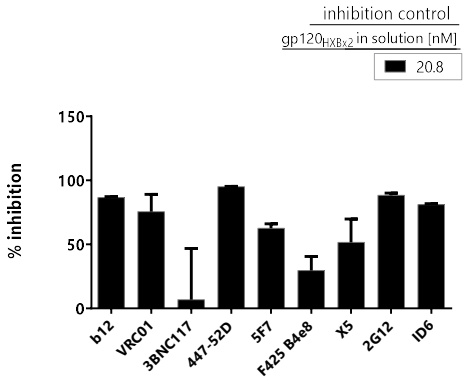


Figure 66: Inhibition of antibody binding to $gp120_{HXBc2}$ by an inhibition control (20.8 nM $gp120_{HXBc2}$), tested in Competitive Assay Format D at one concentration, n=2.

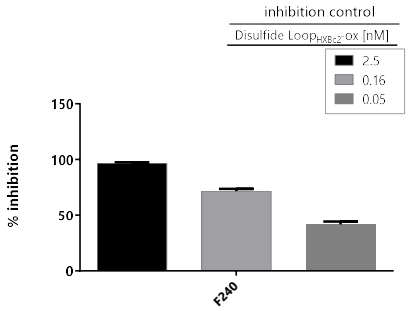


Figure 67: Inhibition of F240 binding to gp41$_{HXBc2}$ by an inhibition control (Disulfide Loop$_{HXBc2}$-ox peptide), tested in competitive Assay Format C at three concentrations (2.5/0.16/0.05 nM), n=2.

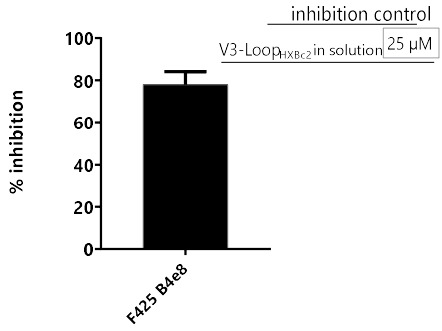


Figure 68: Inhibition of F425 B4e8 - V3-Loop$_{HXBc2}$-ox binding by an inhibition control (V3-Loop$_{HXBc2}$-ox), tested in Competitive Assay Format E at a single concentration (25 µM), n=2.

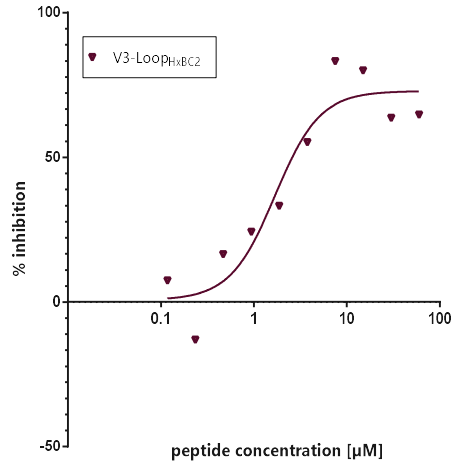


Figure 69: HIV-1_{NL4-3} inhibition by V3-$Loop_{HXBc2}$ peptide. Inhibition curve of V3-$Loop_{HXBc2}$ tested in SEAP Assay.

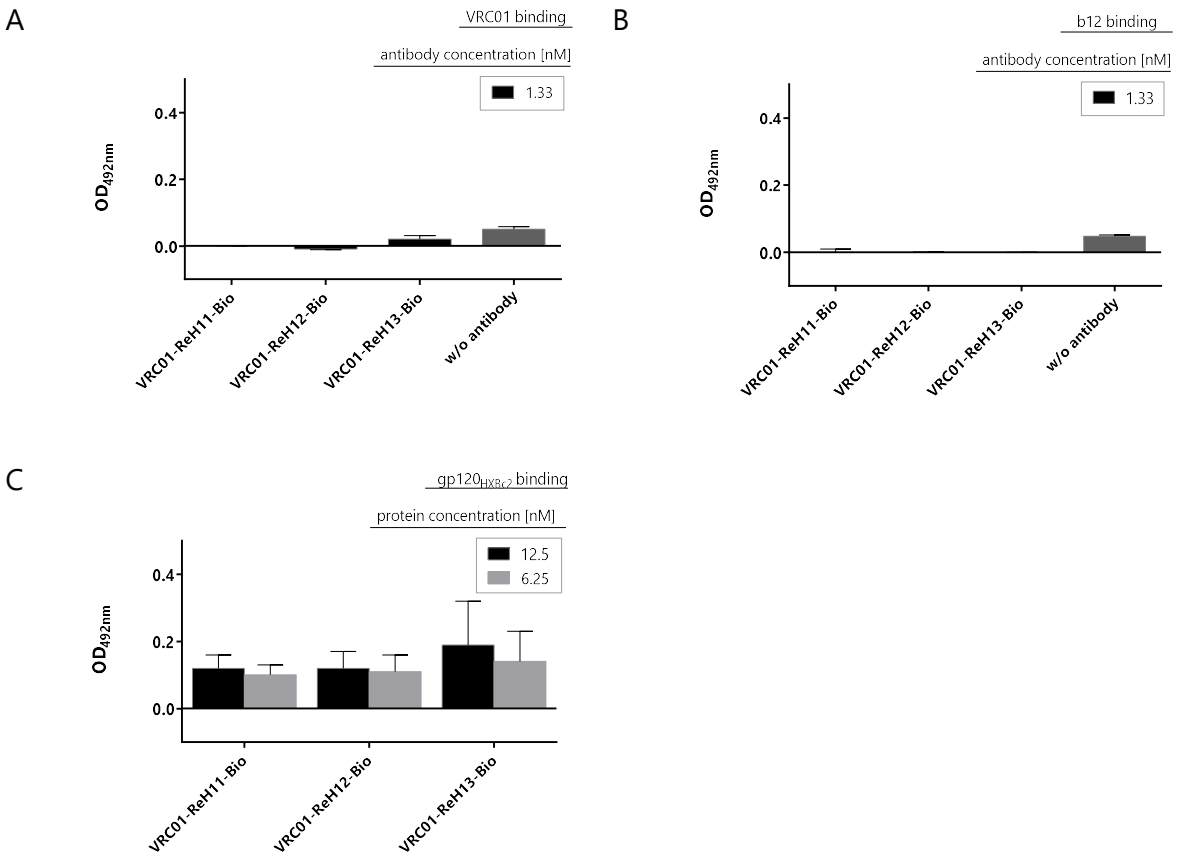


Figure 70: Binding of VRC01-ReH11-Bio, VRC01-ReH12-Bio and VRC01-ReH13-Bio to VRC01, b12 and $gp120_{HXBc2}$, tested in Binding Assay Format B. A: Binding to VRC01 (1.33 nM) B: Binding to b12 (1.33 nM) C: Binding to $gp120_{HXBc2}$ (12.5/6.25 nM).

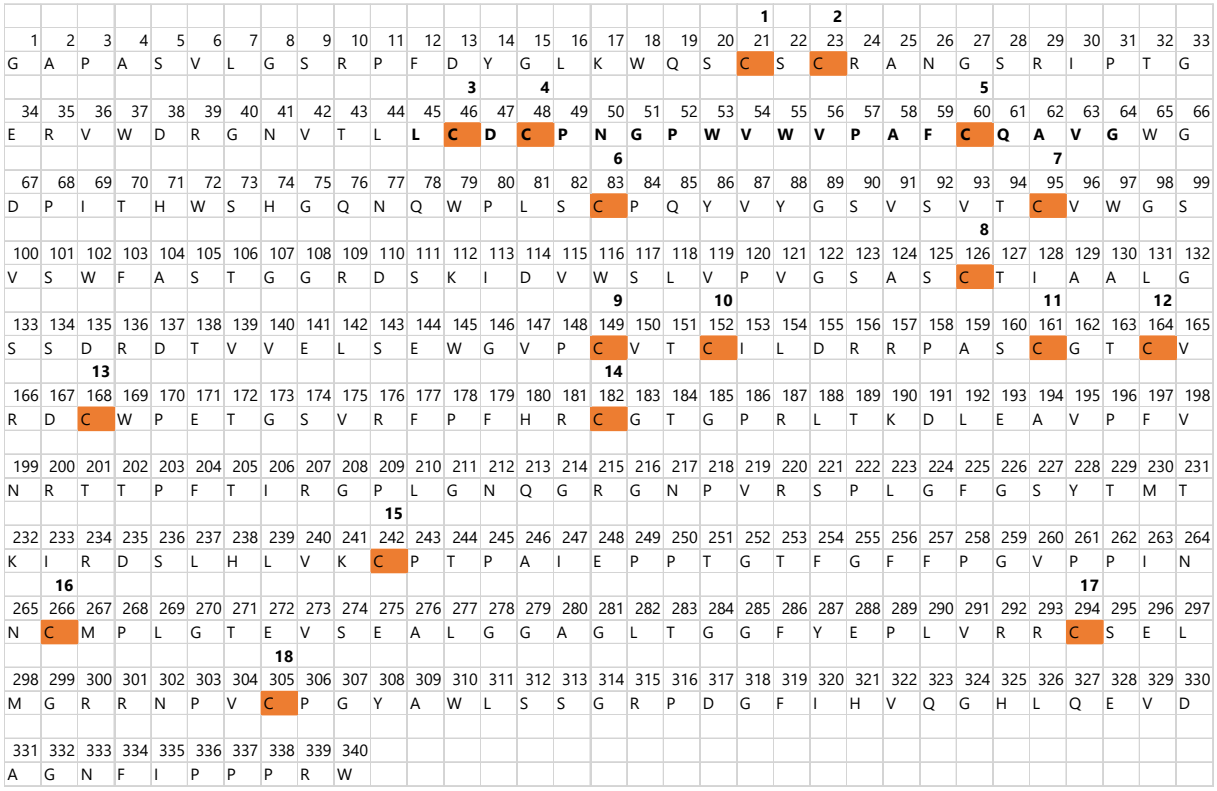


Figure 71: Schematic of the E2 protein sequence AA 1-340 of HPgV-1 Iowa isolate (Genbank Accesion Number AF121950), cysteines are highlighted in orange and additionally numbered (bold), PEGI sequence is depicted in bold letters.

DiANNA webserver 1.1		
Sequence inputSeq Length 340 residues		
Cysteines in this sequence: 18		
Cysteine Class prediction (Half-cystines vs Free cysteines)		
Cysteine	Free cysteine	Half-cystine
21	0.77	0.23
23	0.82	0.18
46	**0.74**	**0.26**
48	**0.81**	**0.19**
60	**0.79**	**0.21**
83	0.45	0.55
95	0.53	0.47
126	0.56	0.44
149	0.80	0.20
152	0.83	0.17
161	0.76	0.24
164	0.58	0.42
168	0.58	0.42
182	0.48	0.52
242	0.63	0.37
266	0.78	0.22
294	0.71	0.29
305	0.71	0.29

Figure 72: Prediction results of a cysteine class prediction (Half-cystines vs Free cysteines) of E2 protein sequence (AA 1-340, Iowa isolate, Genbank Accesion Number AF121950) created with DiANNA webserver © Boston College [132-134].

DiANNA webserver 1.1

Sequence inputSeq Length 340 residues
Cysteines in this sequence: 18
Disulfide Connectivity prediction
Step 1: Running PSI-BLAST with input sequence; click here to see the output
Step 2: Predicting secondary structure using PSIPRED; click here to see the output
Step 3: Disulfide Oxidation State Prediction; click here to see the results
Step 4: Disulfide Bonds Prediction using a trained Neural Network
Disulfide bond scores

Cysteine sequence position	Distance	Bond	Score	Cysteine sequence position	Distance	Bond	Score
21 - 23	2	LKWQSCSCRAN-WQSCSCRANGS	0.01037	83 - 149	66	QWPLSCPQYVY-EWGVPCVTCIL	0.99839
21 - 46	25	LKWQSCSCRAN-NVTLLCDCPNG	0.0104	83 - 152	69	QWPLSCPQYVY-VPCVTCILDRR	0.01062
21 - 48	27	LKWQSCSCRAN-TLLCDCPNGPW	0.01043	83 - 161	78	QWPLSCPQYVY-RRPASCGTCVR	0.01296
21 - 60	39	LKWQSCSCRAN-WVPAFCQAVGW	0.01041	83 - 164	81	QWPLSCPQYVY-ASCGTCVRDCW	0.99916
21 - 83	62	LKWQSCSCRAN-QWPLSCPQYVY	0.0105	83 - 168	85	QWPLSCPQYVY-TCVRDCWPETG	0.02927
21 - 95	74	LKWQSCSCRAN-SVSVTCVWGSV	0.01079	83 - 182	99	QWPLSCPQYVY-FPFHRCGTGPR	0.0125
21 - 126	105	LKWQSCSCRAN-VGSASCTIAAL	0.01073	83 - 242	159	QWPLSCPQYVY-LHLVKCPTPAI	0.75375
21 - 149	128	LKWQSCSCRAN-EWGVPCVTCIL	0.02278	83 - 266	183	QWPLSCPQYVY-PPINNCMPLGT	0.01044
21 - 152	131	LKWQSCSCRAN-VPCVTCILDRR	0.01037	83 - 294	211	QWPLSCPQYVY-PLVRRCSELMG	0.01037
21 - 161	140	LKWQSCSCRAN-RRPASCGTCVR	0.0114	83 - 305	222	QWPLSCPQYVY-RRNPVCPGYAW	0.01598
21 - 164	143	LKWQSCSCRAN-ASCGTCVRDCW	0.98247	95 - 126	31	SVSVTCVWGSV-VGSASCTIAAL	0.01174
21 - 168	147	LKWQSCSCRAN-TCVRDCWPETG	0.01195	95 - 149	54	SVSVTCVWGSV-EWGVPCVTCIL	0.01041
21 - 182	161	LKWQSCSCRAN-FPFHRCGTGPR	0.25913	95 - 152	57	SVSVTCVWGSV-VPCVTCILDRR	0.01039
21 - 242	221	LKWQSCSCRAN-LHLVKCPTPAI	0.01124	95 - 161	66	SVSVTCVWGSV-RRPASCGTCVR	0.07038
21 - 266	245	LKWQSCSCRAN-PPINNCMPLGT	0.14902	95 - 164	69	SVSVTCVWGSV-ASCGTCVRDCW	0.99533
21 - 294	273	LKWQSCSCRAN-PLVRRCSELMG	0.01054	95 - 168	73	SVSVTCVWGSV-TCVRDCWPETG	0.01043
21 - 305	284	LKWQSCSCRAN-RRNPVCPGYAW	0.01217	95 - 182	87	SVSVTCVWGSV-FPFHRCGTGPR	0.01047
23 - 46	23	WQSCSCRANGS-NVTLLCDCPNG	0.14433	95 - 242	147	SVSVTCVWGSV-LHLVKCPTPAI	0.02218
23 - 48	25	WQSCSCRANGS-TLLCDCPNGPW	0.99506	95 - 266	171	SVSVTCVWGSV-PPINNCMPLGT	0.01042
23 - 60	37	WQSCSCRANGS-WVPAFCQAVGW	0.01143	95 - 294	199	SVSVTCVWGSV-PLVRRCSELMG	0.01037
23 - 83	60	WQSCSCRANGS-QWPLSCPQYVY	0.01148	95 - 305	210	SVSVTCVWGSV-RRNPVCPGYAW	0.05618
23 - 95	72	WQSCSCRANGS-SVSVTCVWGSV	0.01632	126 - 149	23	VGSASCTIAAL-EWGVPCVTCIL	0.01613
23 - 126	103	WQSCSCRANGS-VGSASCTIAAL	0.75834	126 - 152	26	VGSASCTIAAL-VPCVTCILDRR	0.01318
23 - 149	126	WQSCSCRANGS-EWGVPCVTCIL	0.9992	126 - 161	35	VGSASCTIAAL-RRPASCGTCVR	0.01485
23 - 152	129	WQSCSCRANGS-VPCVTCILDRR	0.08105	126 - 164	38	VGSASCTIAAL-ASCGTCVRDCW	0.99937
23 - 161	138	WQSCSCRANGS-RRPASCGTCVR	0.86328	126 - 168	42	VGSASCTIAAL-TCVRDCWPETG	0.01076
23 - 164	141	WQSCSCRANGS-ASCGTCVRDCW	0.99973	126 - 182	56	VGSASCTIAAL-FPFHRCGTGPR	0.9977
23 - 168	145	WQSCSCRANGS-TCVRDCWPETG	0.99825	126 - 242	116	VGSASCTIAAL-LHLVKCPTPAI	0.99729
23 - 182	159	WQSCSCRANGS-FPFHRCGTGPR	0.0134	126 - 266	140	VGSASCTIAAL-PPINNCMPLGT	0.01047
23 - 242	219	WQSCSCRANGS-LHLVKCPTPAI	0.78046	126 - 294	168	VGSASCTIAAL-PLVRRCSELMG	0.02222
23 - 266	243	WQSCSCRANGS-PPINNCMPLGT	0.99966	126 - 305	179	VGSASCTIAAL-RRNPVCPGYAW	0.01807
23 - 294	271	WQSCSCRANGS-PLVRRCSELMG	0.0109	149 - 152	3	EWGVPCVTCIL-VPCVTCILDRR	0.03015
23 - 305	282	WQSCSCRANGS-RRNPVCPGYAW	0.9896	149 - 161	12	EWGVPCVTCIL-RRPASCGTCVR	0.94131
46 - 48	2	NVTLLCDCPNG-TLLCDCPNGPW	0.01066	149 - 164	15	EWGVPCVTCIL-ASCGTCVRDCW	0.99974
46 - 60	14	NVTLLCDCPNG-WVPAFCQAVGW	0.0104	149 - 168	19	EWGVPCVTCIL-TCVRDCWPETG	0.99889
46 - 83	37	NVTLLCDCPNG-QWPLSCPQYVY	0.01075	149 - 182	33	EWGVPCVTCIL-FPFHRCGTGPR	0.03234
46 - 95	49	NVTLLCDCPNG-SVSVTCVWGSV	0.01098	149 - 242	93	EWGVPCVTCIL-LHLVKCPTPAI	0.98561
46 - 126	80	NVTLLCDCPNG-VGSASCTIAAL	0.01069	149 - 266	117	EWGVPCVTCIL-PPINNCMPLGT	0.32366
46 - 149	103	NVTLLCDCPNG-EWGVPCVTCIL	0.99977	149 - 294	145	EWGVPCVTCIL-PLVRRCSELMG	0.01075
46 - 152	106	NVTLLCDCPNG-VPCVTCILDRR	0.01046	149 - 305	156	EWGVPCVTCIL-RRNPVCPGYAW	0.98703
46 - 161	115	NVTLLCDCPNG-RRPASCGTCVR	0.84034	152 - 161	9	VPCVTCILDRR-RRPASCGTCVR	0.01322
46 - 164	118	NVTLLCDCPNG-ASCGTCVRDCW	0.02958	152 - 164	12	VPCVTCILDRR-ASCGTCVRDCW	0.07459
46 - 168	122	NVTLLCDCPNG-TCVRDCWPETG	0.01264	152 - 168	16	VPCVTCILDRR-TCVRDCWPETG	0.01044
46 - 182	136	NVTLLCDCPNG-FPFHRCGTGPR	0.01118	152 - 182	30	VPCVTCILDRR-FPFHRCGTGPR	0.98565
46 - 242	196	NVTLLCDCPNG-LHLVKCPTPAI	0.99694	152 - 242	90	VPCVTCILDRR-LHLVKCPTPAI	0.01188
46 - 266	220	NVTLLCDCPNG-PPINNCMPLGT	0.80487	152 - 266	114	VPCVTCILDRR-PPINNCMPLGT	0.035
46 - 294	248	NVTLLCDCPNG-PLVRRCSELMG	0.01042	152 - 294	142	VPCVTCILDRR-PLVRRCSELMG	0.01271
46 - 305	259	NVTLLCDCPNG-RRNPVCPGYAW	0.02098	152 - 305	153	VPCVTCILDRR-RRNPVCPGYAW	0.01077
48 - 60	12	TLLCDCPNGPW-WVPAFCQAVGW	0.01043	161 - 164	3	RRPASCGTCVR-ASCGTCVRDCW	0.94653
48 - 83	35	TLLCDCPNGPW-QWPLSCPQYVY	0.06625	161 - 168	7	RRPASCGTCVR-TCVRDCWPETG	0.01206
48 - 95	47	TLLCDCPNGPW-SVSVTCVWGSV	0.01043	161 - 182	21	RRPASCGTCVR-FPFHRCGTGPR	0.0104
48 - 126	78	TLLCDCPNGPW-VGSASCTIAAL	0.99999	161 - 242	81	RRPASCGTCVR-LHLVKCPTPAI	0.01104
48 - 149	101	TLLCDCPNGPW-EWGVPCVTCIL	0.99758	161 - 266	105	RRPASCGTCVR-PPINNCMPLGT	0.01777
48 - 152	104	TLLCDCPNGPW-VPCVTCILDRR	0.01093	161 - 294	133	RRPASCGTCVR-PLVRRCSELMG	0.01037
48 - 161	113	TLLCDCPNGPW-RRPASCGTCVR	0.99978	161 - 305	144	RRPASCGTCVR-RRNPVCPGYAW	0.0461
48 - 164	116	TLLCDCPNGPW-ASCGTCVRDCW	0.9998	164 - 168	4	ASCGTCVRDCW-TCVRDCWPETG	0.75968
48 - 168	120	TLLCDCPNGPW-TCVRDCWPETG	0.04102	164 - 182	18	ASCGTCVRDCW-FPFHRCGTGPR	0.99806
48 - 182	134	TLLCDCPNGPW-FPFHRCGTGPR	0.1619	164 - 242	78	ASCGTCVRDCW-LHLVKCPTPAI	0.99985
48 - 242	194	TLLCDCPNGPW-LHLVKCPTPAI	0.99989	164 - 266	102	ASCGTCVRDCW-PPINNCMPLGT	0.01313
48 - 266	218	TLLCDCPNGPW-PPINNCMPLGT	0.01593	164 - 294	130	ASCGTCVRDCW-PLVRRCSELMG	0.98403
48 - 294	246	TLLCDCPNGPW-PLVRRCSELMG	0.01043	164 - 305	141	ASCGTCVRDCW-RRNPVCPGYAW	0.99959
48 - 305	257	TLLCDCPNGPW-RRNPVCPGYAW	1	168 - 182	14	TCVRDCWPETG-FPFHRCGTGPR	0.01649
60 - 83	23	WVPAFCQAVGW-QWPLSCPQYVY	0.01043	168 - 242	74	TCVRDCWPETG-LHLVKCPTPAI	0.99815
60 - 95	35	WVPAFCQAVGW-SVSVTCVWGSV	0.01039	168 - 266	98	TCVRDCWPETG-PPINNCMPLGT	0.05346
60 - 126	66	WVPAFCQAVGW-VGSASCTIAAL	0.01434	168 - 294	126	TCVRDCWPETG-PLVRRCSELMG	0.0308
60 - 149	89	WVPAFCQAVGW-EWGVPCVTCIL	0.01082	168 - 305	137	TCVRDCWPETG-RRNPVCPGYAW	0.99594
60 - 152	92	WVPAFCQAVGW-VPCVTCILDRR	0.01041	182 - 242	60	FPFHRCGTGPR-LHLVKCPTPAI	0.01042
60 - 161	101	WVPAFCQAVGW-RRPASCGTCVR	0.69066	182 - 266	84	FPFHRCGTGPR-PPINNCMPLGT	0.29445
60 - 164	104	WVPAFCQAVGW-ASCGTCVRDCW	0.99951	182 - 294	112	FPFHRCGTGPR-PLVRRCSELMG	0.01037
60 - 168	108	WVPAFCQAVGW-TCVRDCWPETG	0.03807	182 - 305	123	FPFHRCGTGPR-RRNPVCPGYAW	0.0106
60 - 182	122	WVPAFCQAVGW-FPFHRCGTGPR	0.01039	242 - 266	24	LHLVKCPTPAI-PPINNCMPLGT	0.01177
60 - 242	182	WVPAFCQAVGW-LHLVKCPTPAI	0.03073	242 - 294	52	LHLVKCPTPAI-PLVRRCSELMG	0.01038
60 - 266	206	WVPAFCQAVGW-PPINNCMPLGT	0.99965	242 - 305	63	LHLVKCPTPAI-RRNPVCPGYAW	0.92387
60 - 294	234	WVPAFCQAVGW-PLVRRCSELMG	0.01043	266 - 294	28	PPINNCMPLGT-PLVRRCSELMG	0.01073
60 - 305	245	WVPAFCQAVGW-RRNPVCPGYAW	0.43817	266 - 305	39	PPINNCMPLGT-RRNPVCPGYAW	0.01127
83 - 95	12	QWPLSCPQYVY-SVSVTCVWGSV	0.01065	294 - 305	11	PLVRRCSELMG-RRNPVCPGYAW	0.99937
83 - 126	43	QWPLSCPQYVY-VGSASCTIAAL	0.02284				

Step 5: Weighted matching

Predicted bonds

21 - 164	LKWQSCSCRAN - ASCGTCVRDCW
23 - 168	WQSCSCRANGS - TCVRDCWPETG
46 - 242	NVTLLCDCPNG - LHLVKCPTPAI
48 - 126	TLLCDCPNGPW - VGSASCTIAAL
60 - 266	WVPAFCQAVGW - PPINNCMPLGT
83 - 149	QWPLSCPQYVY - EWGVPCVTCIL
	SVSVTCVWGSV - RRPASCGTCVR
152 - 182	VPCVTCILDRR - FPFHRCGTGPR
294 - 305	PLVRRCSELMG - RRNPVCPGYAW

Predicted connectivity

1-12, 2-13, 3-15, 4-8, 5-16, 6-9, 7-11, 10-14, 17-18

Figure 73: Results of disulfide connectivity prediction of E2 protein sequence (AA 1-340, Iowa isolate, Genbank Accesion Number AF121950) created with DiANNA webserver © Boston College [132-134].

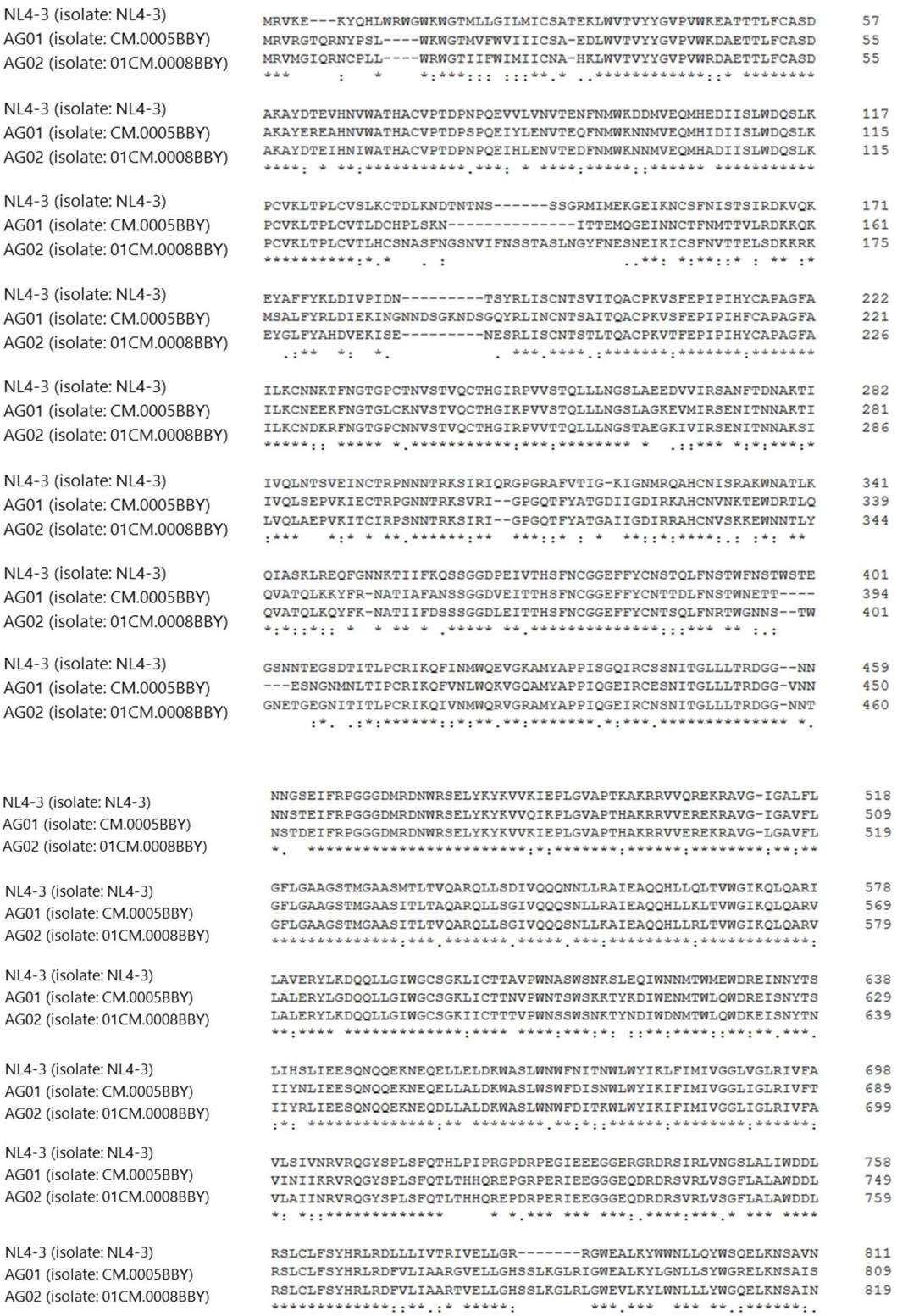


Figure 74: ENV sequence alignment of HIV-1 strains NL4-3, AG01 and AG02. ***** = no difference, **.** = one strain different AA, **:** two strains same AA. Alignment created with Clustal Omega [119].

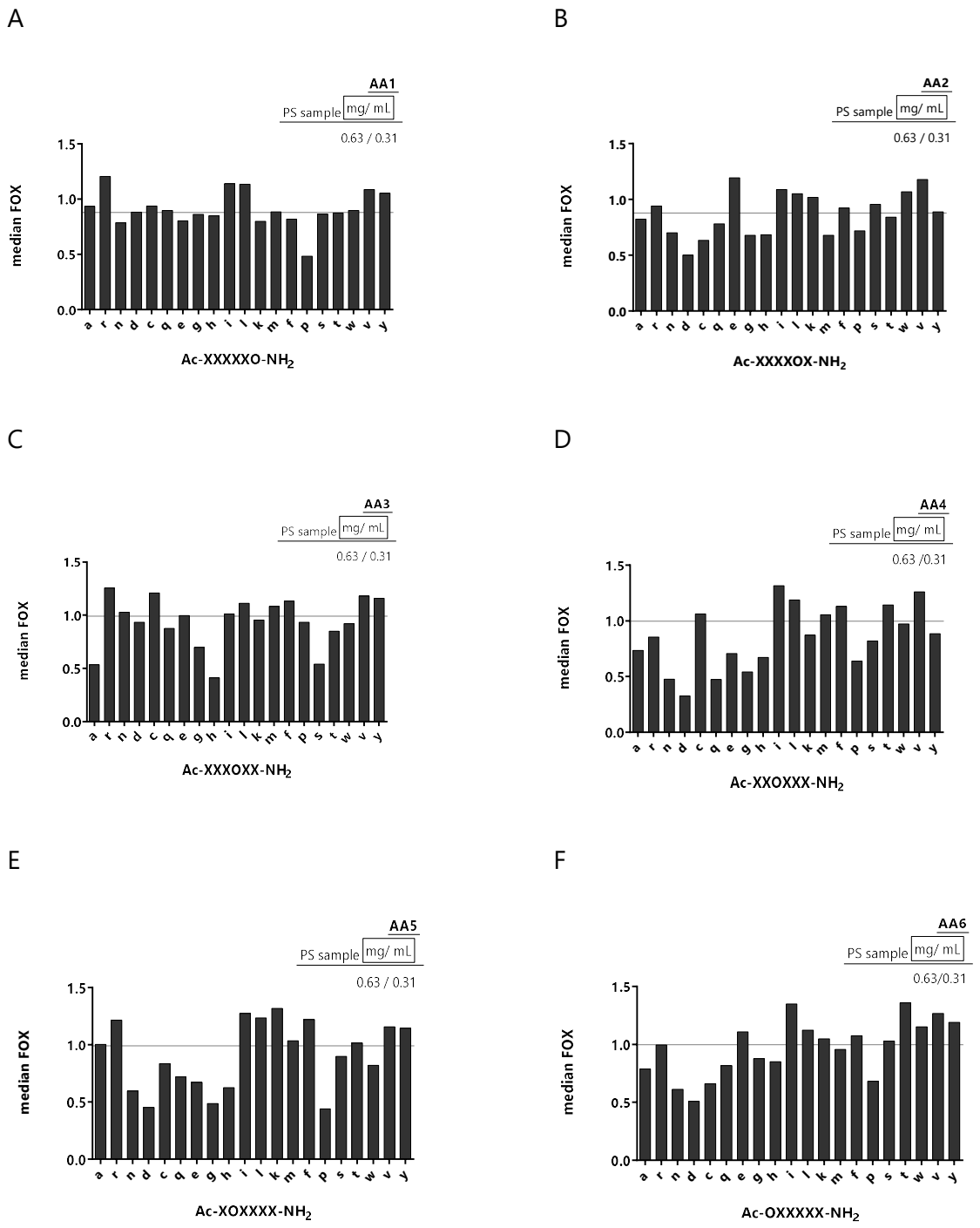


Figure 75: Inhibition of the $g120_{HXBc2}$-b12 interaction by mixtures of PS-SCL TIP 2338 (D-AA hexapeptide library) shown as median of fold over x (FOX), tested in Competitive Assay Format B 0.625/0.312 mg/mL) A) AA position 6 samples B) AA position 5 samples C) AA position 4 samples D) AA position 3 samples E) AA position 2 samples F) AA position 1 samples. A-F: green labeled bars = chosen for deconvolution. O= defines each subsequent position X = mixture of all AA

7.3 List of Figures

7.4 List of Tables

8 Bibliography

1. Jamieson AG, Boutard N, Sabatino D, Lubell WD. Peptide Scanning for Studying Structure-Activity Relationships in Drug Discovery. *Chemical Biology & Drug Design* 2013; 81(1):148-165.
2. Groß A, Hashimoto C, Sticht H, Eichler J. Synthetic Peptides as Protein Mimics. *Frontiers in Bioengineering and Biotechnology* 2015; 3:211.
3. Merrifield RB. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *Journal of the American Chemical Society* 1963; 85(14):2149-2154.
4. Sewald N, Jakubke HD. Peptides: Chemistry and Biology. Wiley; 2015.
5. Fosgerau K, Hoffmann T. Peptide therapeutics: current status and future directions. *Drug Discovery Today* 2015; 20(1):122-128.
6. Eichler J. Peptides as protein binding site mimetics. *Curr Opin Chem Biol* 2008; 12(6):707-713.
7. Robinson JA. Max Bergmann lecture Protein epitope mimetics in the age of structural vaccinology. *Journal of Peptide Science* 2013; 19(3):127-140.
8. Wild C, Greenwell T, Matthews T. A Synthetic Peptide from HIV-1 gp41 Is a Potent Inhibitor of Virus-Mediated Cell—Cell Fusion. *AIDS research and human retroviruses* 1993; 9(11):1051-1053.
9. Matthews T, Salgo M, Greenberg M, Chung J, DeMasi R, Bolognesi D. Enfuvirtide: the first therapy to inhibit the entry of HIV-1 into host CD4 lymphocytes. *Nat Rev Drug Discov* 2004; 3(3):215-225.
10. Haußner C, Damm D, Nirschl S, Rohrhofer A, Schmidt B, Eichler J. Peptide Paratope Mimics of the Broadly Neutralizing HIV-1 Antibody b12. *ChemBioChem* 2017; 18(7):647-653.
11. Groß A, Brox R, Damm D, Tschammer N, Schmidt B, Eichler J. Ligand selectivity of a synthetic CXCR4 mimetic peptide. *Bioorganic & medicinal chemistry* 2015; 23(14):4050-4055.
12. Mobius K, Durr R, Haussner C, Dietrich U, Eichler J. A functionally selective synthetic mimic of the HIV-1 co-receptor CXCR4. *Chemistry* 2012; 18(27):8292-8295.
13. Gross A, Mobius K, Haussner C, Donhauser N, Schmidt B, Eichler J. Mimicking Protein-Protein Interactions through Peptide-Peptide Interactions: HIV-1 gp120 and CXCR4. *Front Immunol* 2013; 4:257.
14. UNAIDS. UNAIDS Data 2017. 2017. *http://www.unaids.org/en/resources/documents/2017/2017_data_book; December 16, 2017*
15. UNAIDS. Global AIDS Update 2016. 2017. *www.who.int/entity/hiv/pub/arv/global-AIDS-update-2016_en.pdf; December 16. 2017*
16. Arts EJ, Hazuda DJ. HIV-1 Antiretroviral Drug Therapy. *Cold Spring Harbor perspectives in medicine* 2012; 2(4):a007161.
17. Reeves JD, Doms RW. Human immunodeficiency virus type 2. *Journal of General Virology* 2002; 83(6):1253-1265.
18. Sharp PM, Hahn BH. Origins of HIV and the AIDS Pandemic. *Cold Spring Harbor Perspectives in Medicine:* 2011; 1(1):a006841.
19. Jakobsdottir GM, Iliopoulou M, Nolan R, Alvarez L, Compton AA, Padilla-Parra S. On the Whereabouts of HIV-1 Cellular Entry and Its Fusion Ports. *Trends in Molecular Medicine* 2017; 23(10):932-944.
20. Gallo S, Finnegan C, Viard M, Raviv Y, Dimitrov A, Rawat S, et al. The HIV Env-mediated fusion reaction. *Biochimica et biophysica acta* 2003; 1614(1):36 - 50.
21. Tilton JC, Doms RW. Entry inhibitors in the treatment of HIV-1 infection. *Antiviral research* 2010; 85(1):91-100.
22. Eissmann K, Mueller S, Sticht H, Jung S, Zou P, Jiang S, et al. HIV-1 fusion is blocked through binding of GB Virus C E2D peptides to the HIV-1 gp41 disulfide loop. *PloS one* 2013; 8(1):e54452.
23. Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* 1998; 393:648.
24. Wilen CB, Tilton JC, Doms RW. HIV: cell binding and entry. *Cold Spring Harbor perspectives in medicine* 2012; 2(8).

25. Hwang S, Boyle T, Lyerly H, Cullen B. Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1. *Science* 1991; 253(5015):71-74.
26. Hartley O, Klasse PJ, Sattentau QJ, Moore JP. V3: HIV's Switch-Hitter. *AIDS research and human retroviruses* 2005; 21(2):171-189.
27. Burton DR, Montefiori DC. The antibody response in HIV-1 infection. *Aids* 1997; 11 Suppl A:S87-98.
28. Mascola JR, Haynes BF. HIV-1 neutralizing antibodies: understanding nature's pathways. *Immunological Reviews* 2013; 254(1):225-244.
29. Kwong Peter D, Mascola John R. Human Antibodies that Neutralize HIV-1: Identification, Structures, and B Cell Ontogenies. *Immunity* 2012; 37(3):412-425.
30. Mascola JR, Montefiori DC. The Role of Antibodies in HIV Vaccines. *Annual Review of Immunology* 2010; 28(1):413-444.
31. McElrath MJ, Haynes BF. Induction of Immunity to Human Immunodeficiency Virus Type-1 By Vaccination. *Immunity* 2010; 33(4):542-554.
32. Bradley T, Pollara J, Santra S, Vandergrift N, Pittala S, Bailey-Kellogg C, et al. Pentavalent HIV-1 vaccine protects against simian-human immunodeficiency virus challenge. *Nat Commun* 2017; 8:15711.
33. Tomaras GD, Plotkin SA. Complex immune correlates of protection in HIV-1 vaccine efficacy trials. *Immunological Reviews* 2017; 275(1):245-261.
34. Karnasuta C, Akapirat S, Madnote S, Savadsuk H, Puangkaew J, Rittiroongrad S, et al. Comparison of Antibody Responses Induced by RV144, VAX003, and VAX004 Vaccination Regimens. *AIDS research and human retroviruses* 2017; 33(5):410-423.
35. McCoy LE, Burton DR. Identification and specificity of broadly neutralizing antibodies against HIV. *Immunological Reviews* 2017; 275(1):11-20.
36. Kwong PD, Doyle ML, Casper DJ, Cicala C, Leavitt SA, Majeed S, et al. HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites. *Nature* 2002; 420(6916):678-682.
37. Karlsson Hedestam GB, Fouchier RA, Phogat S, Burton DR, Sodroski J, Wyatt RT. The challenges of eliciting neutralizing antibodies to HIV-1 and to influenza virus. *Nat Rev Microbiol* 2008; 6(2):143-155.
38. Pancera M, Zhou T, Druz A, Georgiev IS, Soto C, Gorman J, et al. Structure and immune recognition of trimeric pre-fusion HIV-1 Env. *Nature* 2014.
39. Kwong PD, Mascola JR, Nabel GJ. Broadly neutralizing antibodies and the search for an HIV-1 vaccine: the end of the beginning. *Nat Rev Immunol* 2013; 13(9):693-701.
40. Chen J, Frey G, Peng H, Rits-Volloch S, Garrity J, Seaman MS, et al. Mechanism of HIV-1 Neutralization by Antibodies Targeting a Membrane-Proximal Region of gp41. *Journal of virology* 2014; 88(2):1249-1258.
41. Brunel FM, Zwick MB, Cardoso RMF, Nelson JD, Wilson IA, Burton DR, et al. Structure-Function Analysis of the Epitope for 4E10, a Broadly Neutralizing Human Immunodeficiency Virus Type 1 Antibody. *Journal of virology* 2006; 80(4):1680-1687.
42. Irimia A, Sarkar A, Stanfield Robyn L, Wilson Ian A. Crystallographic Identification of Lipid as an Integral Component of the Epitope of HIV Broadly Neutralizing Antibody 4E10. *Immunity*; 44(1):21-31.
43. Scanlan CN, Pantophlet R, Wormald MR, Ollmann Saphire E, Stanfield R, Wilson IA, et al. The Broadly Neutralizing Anti-Human Immunodeficiency Virus Type 1 Antibody 2G12 Recognizes a Cluster of α1→2 Mannose Residues on the Outer Face of gp120. *Journal of virology* 2002; 76(14):7306-7321.
44. McLellan JS, Pancera M, Carrico C, Gorman J, Julien J-P, Khayat R, et al. Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature* 2011; 480:336.
45. Pancera M, McLellan JS, Wu X, Zhu J, Changela A, Schmidt SD, 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. *Journal of virology* 2010; 84(16):8098-8110.
46. Pancera M, Shahzad-ul-Hussan S, Doria-Rose NA, McLellan JS, Bailer RT, Dai K, et al. Structural basis for diverse N-glycan recognition by HIV-1–neutralizing V1–V2–directed antibody PG16. *Nature Structural &Amp; Molecular Biology* 2013; 20:804.

47. Zhou T, Xu L, Dey B, Hessell AJ, Van Ryk D, Xiang S-H, et al. Structural definition of a conserved neutralization epitope on HIV-1 gp120. *Nature* 2007; 445(7129):732-737.
48. Zhou T, Georgiev I, Wu X, Yang ZY, Dai K, Finzi A, et al. Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science* 2010; 329(5993):811-817.
49. Wu X, Yang ZY, Li Y, Hogerkorp CM, Schief WR, Seaman MS, et al. Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* 2010; 329(5993):856-861.
50. Wu X, Zhou T, Zhu J, Zhang B, Georgiev I, Wang C, et al. Focused Evolution of HIV-1 Neutralizing Antibodies Revealed by Structures and Deep Sequencing. *Science* 2011; 333(6049):1593.
51. Zhou T, Zhu J, Wu X, Moquin S, Zhang B, Acharya P, et al. Multidonor Analysis Reveals Structural Elements, Genetic Determinants, and Maturation Pathway for HIV-1 Neutralization by VRC01-Class Antibodies. *Immunity*; 39(2):245-258.
52. ClinicalTrials.gov. VRC 601: A Phase I, Open-Label, Dose-Escalation Study of the Safety and Pharmacokinetics of a Human Monoclonal Antibody, VRC HIVMAB060-00-AB (VRC01), With Broad HIV-1 Neutralizing Activity, Administered Intravenously or Subcutaneously to HIV-Infected Adults. *https://clinicaltrials.gov/ct2/show/NCT01950325; December 11, 2017*
53. Diseases NIoAaI. NIH Launches Large Clinical Trials of Antibody-Based HIV Prevention. *https://www.niaid.nih.gov/news-events/nih-launches-large-clinical-trials-antibody-based-hiv-prevention; December 11, 2017*
54. Huang J, Kang Byong H, Ishida E, Zhou T, Griesman T, Sheng Z, et al. Identification of a CD4-Binding-Site Antibody to HIV that Evolved Near-Pan Neutralization Breadth. *Immunity* 2016; 45(5):1108-1121.
55. Huang J, Kang BH, Pancera M, Lee JH, Tong T, Feng Y, et al. Broad and potent HIV-1 neutralization by a human antibody that binds the gp41-gp120 interface. *Nature* 2014; 515(7525):138-142.
56. Xu L, Pegu A, Rao E, Doria-Rose N, Beninga J, McKee K, et al. Trispecific broadly neutralizing HIV antibodies mediate potent SHIV protection in macaques. *Science* 2017; 358(6359):85.
57. Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, et al. HIV-1 Nomenclature Proposal. *Science* 2000; 288(5463):55.
58. Buonaguro L, Tornesello ML, Buonaguro FM. Human Immunodeficiency Virus Type 1 Subtype Distribution in the Worldwide Epidemic: Pathogenetic and Therapeutic Implications. *Journal of virology* 2007; 81(19):10209-10219.
59. Linnen J, Wages J, Zhang-Keck Z-Y, Fry KE, Krawczynski KZ, Alter H, et al. Molecular Cloning and Disease Association of Hepatitis G Virus: A Transfusion-Transmissible Agent. *Science* 1996; 271(5248):505.
60. Simons JN, Leary TP, Dawson GJ, Pilot-Matias TJ, Muerhoff AS, Schlauder GG, et al. Isolation of novel virus-like sequences associated with human hepatitis. *Nature Medicine* 1995; 1:564.
61. Leary TP, Muerhoff AS, Simons JN, Pilot-Matias TJ, Erker JC, Chalmers ML, et al. Sequence and genomic organization of GBV-C: a novel member of the flaviviridae associated with human non-A-E hepatitis. *Journal of medical virology* 1996; 48(1):60-67.
62. Adams MJ, Lefkowitz EJ, King AMQ, Bamford DH, Breitbart M, Davison AJ, et al. Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2015). *Archives of Virology* 2015; 160(7):1837-1850.
63. Alter HJ, Nakatsuji Y, Melpolder J, Wages J, Wesley R, Shih JW-K, et al. The Incidence of Transfusion-Associated Hepatitis G Virus Infection and Its Relation to Liver Disease. *New England Journal of Medicine* 1997; 336(11):747-754.
64. Stapleton JT. GB Virus Type C/Hepatitis G Virus. *Semin Liver Dis* 2003; 23(02):137-148.
65. Chivero ET, Stapleton JT. Tropism of human pegivirus (formerly known as GB virus C/hepatitis G virus) and host immunomodulation: insights into a highly successful viral infection. *Journal of General Virology* 2015; 96(7):1521-1532.
66. Mohr EL, Stapleton JT. GB virus type C interactions with HIV: the role of envelope glycoproteins. *Journal of viral hepatitis* 2009; 16(11):757-768.
67. Tucker TJ, Smuts HE. GBV-C/HGV genotypes: Proposed nomenclature for genotypes 1–5. *Journal of medical virology* 2000; 62(1):82-83.

68. Smith DB, Basaras M, Frost S, Haydon D, Cuceanu N, Prescott L, et al. Phylogenetic analysis of GBV-C/hepatitis G virus. *Journal of General Virology* 2000; 81(3):769-780.
69. Liu H-F, Teng C-W, Fukuda Y, Nakano I, Hayashi K, Takamatsu J, et al. A novel subtype of GB virus C/hepatitis G virus genotype 1 detected uniquely in patients with hemophilia in Japan. *Journal of medical virology* 2003; 71(3):385-390.
70. Muerhoff AS, Smith DB, Leary TP, Erker JC, Desai SM, Mushahwar IK. Identification of GB virus C variants by phylogenetic analysis of 5'-untranslated and coding region sequences. *Journal of virology* 1997; 71(9):6501-6508.
71. Feng Y, Zhao W, Feng Y, Dai J, Li Z, Zhang X, et al. A Novel Genotype of GB Virus C: Its Identification and Predominance among Injecting Drug Users in Yunnan, China. *PloS one* 2011; 6(10):e21151.
72. Muerhoff AS, Dawson GJ, Desai SM. A previously unrecognized sixth genotype of GB virus C revealed by analysis of 5′-untranslated region sequences. *Journal of medical virology* 2006; 78(1):105-111.
73. Mison L, Hyland C, Poidinger M, Borthwick I, Faoagali J, Aeno U, et al. Hepatitis G virus genotypes in Australia, Papua New Guinea and the Solomon Islands: A possible new Pacific type identified. *Journal of Gastroenterology and Hepatology* 2000; 15(8):952-956.
74. Simmonds P, Smith DB. Structural Constraints on RNA Virus Evolution. *Journal of virology* 1999; 73(7):5787-5794.
75. Giret MT, Kallas EG. GBV-C: state of the art and future prospects. *Current HIV/AIDS reports* 2012; 9(1):26-33.
76. Wagner RR. Viral interference. Some considerations of basic mechanisms and their potential relationship to host resistance. *Bacteriological Reviews* 1960; 24(1):151-166.
77. Xiang J, George SL, Wünschmann S, Chang Q, Klinzman D, Stapleton JT. Inhibition of HIV-1 replication by GB virus C infection through increases in RANTES, MIP-lα, MIP-1β, and SDF-1. *The Lancet* 2004; 363(9426):2040-2046.
78. Xiang J, McLinden JH, Chang Q, Jordan EL, Stapleton JT. Characterization of a peptide domain within the GB virus C NS5A phosphoprotein that inhibits HIV replication. *PloS one* 2008; 3(7):e2580.
79. Jung S, Eichenmuller M, Donhauser N, Neipel F, Engel AM, Hess G, et al. HIV entry inhibition by the envelope 2 glycoprotein of GB virus C. *Aids* 2007; 21(5):645-647.
80. Kaufman TM, McLinden JH, Xiang J, Engel AM, Stapleton JT. The GBV-C envelope glycoprotein E2 does not interact specifically with CD81. *Aids* 2007; 21(8):1045-1048.
81. McLinden JH, Kaufman TM, Xiang J, Chang Q, Klinzman D, Engel AM, et al. Characterization of an Immunodominant Antigenic Site on GB Virus C Glycoprotein E2 That Is Involved in Cell Binding. *Journal of virology* 2006; 80(24):12131-12140.
82. Chivero ET, Bhattarai N, Rydze RT, Winters MA, Holodniy M, Stapleton JT. Human pegivirus RNA is found in multiple blood mononuclear cells in vivo and serum-derived viral RNA-containing particles are infectious in vitro. *The Journal of General Virology* 2014; 95(Pt 6):1307-1319.
83. Schwarze-Zander C, Neibecker M, Othman S, Tural C, Clotet B, Blackard JT, et al. GB Virus C (GBV-C) coinfection in advanced HIV disease is associated with low CCR5 and CXCR4 surface expression on CD4(+) T cells. *Antiviral therapy* 2010; 15(5):745-752.
84. Jung S. GB Virus C Proteins interfere with the Replication of HIV-1. 2010. *urn:nbn:de:bvb:29-opus-21385; January 10, 2018*
85. Koedel Y, Eissmann K, Wend H, Fleckenstein B, Reil H. Peptides derived from a distinct region of GB virus C glycoprotein E2 mediate strain-specific HIV-1 entry inhibition. *Journal of virology* 2011; 85(14):7037-7047.
86. Herrera E, Tenckhoff S, Gomara MJ, Galatola R, Bleda MJ, Gil C, et al. Effect of synthetic peptides belonging to E2 envelope protein of GB virus C on human immunodeficiency virus type 1 infection. *Journal of medicinal chemistry* 2010; 53(16):6054-6063.
87. Chivero ET, Bhattarai N, McLinden JH, Xiang J, Stapleton JT. Human Pegivirus (HPgV; formerly known as GBV-C) inhibits IL-12 dependent natural killer cell function. *Virology* 2015; 485(Supplement C):116-127.

88. Gómara MJ, Sánchez-Merino V, Paús A, Merino-Mansilla A, Gatell JM, Yuste E, et al. Definition of an 18-mer Synthetic Peptide Derived from the GB virus C E1 Protein as a New HIV-1 Entry Inhibitor. *Biochimica et Biophysica Acta (BBA) - General Subjects* 2016; 1860(6):1139-1148.
89. Liu R, Li X, Lam KS. Combinatorial chemistry in drug discovery. *Current Opinion in Chemical Biology* 2017; 38(Supplement C):117-126.
90. Houghten RA. General method for the rapid solid-phase synthesis of large numbers of peptides: specificity of antigen-antibody interaction at the level of individual amino acids. *Proceedings of the National Academy of Sciences* 1985; 82(15):5131-5135.
91. Houghten RA, Pinilla C, Blondelle SE, Appel JR, Dooley CT, Cuervo JH. Generation and use of synthetic peptide combinatorial libraries for basic research and drug discovery. *Nature* 1991; 354:84.
92. Pinilla C, Appel J, Blondelle S, Dooley C, Dörner B, Eichler J, et al. A review of the utility of soluble peptide combinatorial libraries. *Biopolymers* 1995; 37(3):221-240.
93. Pinilla C, Appel JR, Borras E, Houghten RA. Advances in the use of synthetic combinatorial chemistry: Mixture-based libraries. *Nat Med* 2003; 9(1):118-122.
94. Eichler J. Synthetic Peptide Arrays and Peptide Combinatorial Libraries for the Exploration of Protein-Ligand Interactions and the Design of Protein Inhibitors. *Combinatorial Chemistry & High Throughput Screening* 2005; 8(2):135-143.
95. Houghten RA, Pinilla C, Giulianotti MA, Appel JR, Dooley CT, Nefzi A, et al. Strategies for the Use of Mixture-Based Synthetic Combinatorial Libraries: Scaffold Ranking, Direct Testing In Vivo, and Enhanced Deconvolution by Computational Methods. *Journal of Combinatorial Chemistry* 2008; 10(1):3-19.
96. Hatefi Y, Hanstein WG. Solubilization of particulate proteins and nonelectrolytes by chaotropic agents. *Proceedings of the National Academy of Sciences of the United States of America* 1969; 62(4):1129-1136.
97. Kaiser E, Colescott RL, Bossinger CD, Cook PI. Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Analytical biochemistry* 1970; 34(2):595-598.
98. Ellman GL. Tissue sulfhydryl groups. *Archives of biochemistry and biophysics* 1959; 82(1):70-77.
99. Ellman G, Lysko H. A precise method for the determination of whole blood and plasma sulfhydryl groups. *Analytical biochemistry* 1979; 93(1):98-102.
100. Biancalana M, Koide S. Molecular Mechanism of Thioflavin-T Binding to Amyloid Fibrils. *Biochimica et biophysica acta* 2010; 1804(7):1405-1412.
101. Khurana R, Coleman C, Ionescu-Zanetti C, Carter SA, Krishna V, Grover RK, et al. Mechanism of thioflavin T binding to amyloid fibrils. *Journal of Structural Biology* 2005; 151(3):229-238.
102. Hudson SA, Ecroyd H, Kee TW, Carver JA. The thioflavin T fluorescence assay for amyloid fibril detection can be biased by the presence of exogenous compounds. *The FEBS journal* 2009; 276(20):5960-5972.
103. Lengauer T, Sander O, Sierra S, Thielen A, Kaiser R. Bioinformatics prediction of HIV coreceptor usage. *Nat Biotechnol* 2007; 25(12):1407-1410.
104. Yi MC, Khosla C. Thiol--Disulfide Exchange Reactions in the Mammalian Extracellular Environment. *Annual review of chemical and biomolecular engineering* 2016; 7:197-222.
105. Jacob C, Giles GI, Giles NM, Sies H. Sulfur and Selenium: The Role of Oxidation State in Protein Structure and Function. *Angewandte Chemie International Edition* 2003; 42(39):4742-4758.
106. Mayer JP, Zhang F, DiMarchi RD. Insulin structure and function. *Peptide Science* 2007; 88(5):687-713.
107. de Araujo AD, Herzig V, Windley MJ, Dziemborowicz S, Mobli M, Nicholson GM, et al. Do Vicinal Disulfide Bridges Mediate Functionally Important Redox Transformations in Proteins? *Antioxidants & redox signaling* 2013; 19(16):1976-1980.
108. Lipton SH, Bodwell CE. Oxidation of amino acids by dimethyl sulfoxide. *Journal of agricultural and food chemistry* 1973; 21(2):235-237.
109. Champagne K, Shishido A, Root MJ. Interactions of HIV-1 Inhibitory Peptide T20 with the gp41 N-HR Coiled Coil. *The Journal of biological chemistry* 2009; 284(6):3619-3627.

110. Peisajovich SG, Gallo SA, Blumenthal R, Shai Y. C-terminal Octylation Rescues an Inactive T20 Mutant: Implications for the mechanism of hiv/simian immunodeficiency virus-induced membrane fusion. *Journal of Biological Chemistry* 2003; 278(23):21012-21017.
111. Lawless MK, Barney S, Guthrie KI, Bucy TB, Petteway SR, Merutka G. HIV-1 Membrane Fusion Mechanism: Structural Studies of the Interactions between Biologically-Active Peptides from gp41. *Biochemistry* 1996; 35(42):13697-13708.
112. Ncbi Resource Coordinators. Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Research* 2017; 45:D12-D17.
113. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 1997; 25(17):3389-3402.
114. Ruegamer T, Hoffmann R, Rohrhofer A, Audebert F, Salzberger B, Korn K, et al. Inhibition of HIV-1 infection by HPgV-1 derived peptides is affected by HPgV-1 genotype and HIV-1 coreceptor tropism. *Aids* 2018; Publish Ahead of Print.
115. Foisy MT, A. Selected Properties of Enfuvirtide.
hivclinic.ca/main/drugs_properties_files/FI%20enfuvirtide.pdf; January 3, 2018
116. Fuzeon : EPAR - Product Information.
http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/000514/human_med_000802.jsp&mid=WC0b01ac058001d124; January 3, 2018
117. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* 2011; 7(1).
118. Huerta-Cepas J, Serra F, Bork P. ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Molecular Biology and Evolution* 2016; 33(6):1635-1638.
119. McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, et al. Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Research* 2013; 41(W1):W597-W600.
120. Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, et al. A new bioinformatics analysis tools framework at EMBL–EBI. *Nucleic Acids Research* 2010; 38(suppl_2):W695-W699.
121. Pantophlet R, Aguilar-Sino RO, Wrin T, Cavacini LA, Burton DR. Analysis of the neutralization breadth of the anti-V3 antibody F425-B4e8 and re-assessment of its epitope fine specificity by scanning mutagenesis. *Virology* 2007; 364(2):441-453.
122. Kaufmann SHE. Basiswissen Immunologie. Springer Berlin Heidelberg; 2013.
123. Geysen HM, Rodda SJ, Mason TJ. A priori delineation of a peptide which mimics a discontinuous antigenic determinant. *Molecular Immunology* 1986; 23(7):709-715.
124. Cavacini LA, Emes CL, Wisnewski AV, Power J, Lewis G, Montefiori D, et al. Functional and Molecular Characterization of Human Monoclonal Antibody Reactive with the Immunodominant Region of HIV Type 1 Glycoprotein 41. *AIDS research and human retroviruses* 1998; 14(14):1271-1280.
125. Groß A, Rödel K, Kneidl B, Donhauser N, Mössl M, Lump E. Enhancement and induction of HIV-1 infection through an assembled peptide derived from the CD4 binding site of gp120. *Chem Bio Chem* 2015; 16.
126. Stapleton JT, Foung S, Muerhoff AS, Bukh J, Simmonds P. The GB viruses: a review and proposed classification of GBV-A, GBV-C (HGV), and GBV-D in genus Pegivirus within the family Flaviviridae. *The Journal of General Virology* 2011; 92(Pt 2):233-246.
127. Rost B, Sander C. Prediction of Protein Secondary Structure at Better than 70% Accuracy. *Journal of molecular biology* 1993; 232(2):584-599.
128. Rost B, Yachdav G, Liu J. The PredictProtein server. *Nucleic Acids Research* 2004; 32(Web Server issue):W321-W326.
129. Schlessinger A, Rost B. Protein flexibility and rigidity predicted from sequence. *Proteins: Structure, Function, and Bioinformatics* 2005; 61(1):115-126.
130. Schlessinger A, Punta M, Yachdav G, Kajan L, Rost B. Improved Disorder Prediction by Combination of Orthogonal Approaches. *PloS one* 2009; 4(2):e4433.

131. Trivedi MV, Laurence JS, Siahaan TJ. The role of thiols and disulfides in protein chemical and physical stability. *Current protein & peptide science* 2009; 10(6):614-625.
132. Ferrè F, Clote P. DiANNA: a web server for disulfide connectivity prediction. *Nucleic Acids Research* 2005; 33(Web Server issue):W230-W232.
133. Ferrè F, Clote P. Disulfide connectivity prediction using secondary structure information and diresidue frequencies. *Bioinformatics* 2005; 21(10):2336-2346.
134. Ferrè F, Clote P. DiANNA 1.1: an extension of the DiANNA web server for ternary cysteine classification. *Nucleic Acids Research* 2006; 34(Web Server issue):W182-W185.
135. Wouters MA, Fan SW, Haworth NL. Disulfides as Redox Switches: From Molecular Mechanisms to Functional Significance. *Antioxidants & redox signaling* 2009; 12(1):53-91.
136. Xiang J, McLinden JH, Kaufman TM, Mohr EL, Bhattarai N, Chang Q, et al. Characterization of a peptide domain within the GB virus C envelope glycoprotein (E2) that inhibits HIV replication. *Virology* 2012; 430(1):53-62.
137. Caskey M, Schoofs T, Gruell H, Settler A, Karagounis T, Kreider EF, et al. Antibody 10-1074 suppresses viremia in HIV-1-infected individuals. *Nature Medicine* 2017; 23:185.
138. Baldwin C, Sanders R, Deng Y, Jurriaans S, Lange J, Lu M, et al. Emergence of a drug-dependent human immunodeficiency virus type 1 variant during therapy with the T20 fusion inhibitor. *Journal of virology* 2004; 78(22):12428 - 12437.
139. Los Alamos HIV databases. *http://www.hiv.lanl.gov/. 04 March, 2018*
140. Marshall RD. The nature and metabolism of the carbohydrate-peptide linkages of glycoproteins. *Biochem Soc Symp* 1974; (40):17-26.
141. Huang C-c, Tang M, Zhang M-Y, Majeed S, Montabana E, Stanfield RL, et al. Structure of a V3-Containing HIV-1 gp120 Core. *Science (New York, NY)* 2005; 310(5750):1025-1028.
142. Gorny MK, Revesz K, Williams C, Volsky B, Louder MK, Anyangwe CA, et al. The V3 Loop Is Accessible on the Surface of Most Human Immunodeficiency Virus Type 1 Primary Isolates and Serves as a Neutralization Epitope. *Journal of virology* 2004; 78(5):2394-2404.
143. Huang CC, Venturi M, Majeed S, Moore MJ, Phogat S, Zhang MY, et al. Structural basis of tyrosine sulfation and VH-gene usage in antibodies that recognize the HIV type 1 coreceptor-binding site on gp120. *Proceedings of the National Academy of Sciences of the United States of America* 2004; 101(9):2706-2711.
144. Stanfield RL, Gorny MK, Williams C, Zolla-Pazner S, Wilson IA. Structural Rationale for the Broad Neutralization of HIV-1 by Human Monoclonal Antibody 447-52D. *Structure* 2004; 12(2):193-204.
145. Bell CH, Pantophlet R, Schiefner A, Cavacini LA, Stanfield RL, Burton DR, et al. Structure of antibody F425-B4e8 in complex with a V3 peptide reveals a new binding mode for HIV-1 neutralization. *Journal of molecular biology* 2008; 375(4):969-978.
146. Kumar S, Nussinov R. Close-Range Electrostatic Interactions in Proteins. *ChemBioChem* 2002; 3(7):604-617.
147. Sharon M, Kessler N, Levy R, Zolla-Pazner S, Görlach M, Anglister J. Alternative Conformations of HIV-1 V3 Loops Mimic β Hairpins in Chemokines, Suggesting a Mechanism for Coreceptor Selectivity. *Structure* 2003; 11(2):225-236.
148. Rosen O, Sharon M, Quadt-Akabayov SR, Anglister J. Molecular switch for alternative conformations of the HIV-1 V3 region: Implications for phenotype conversion. *Proceedings of the National Academy of Sciences of the United States of America* 2006; 103(38):13950-13955.
149. Gnanakaran S, Daniels MG, Bhattacharya T, Lapedes AS, Sethi A, Li M, et al. Genetic Signatures in the Envelope Glycoproteins of HIV-1 that Associate with Broadly Neutralizing Antibodies. *PLoS Computational Biology* 2010; 6(10):e1000955.
150. Georgiev IS, Joyce MG, Yang Y, Sastry M, Zhang B, Baxa U, et al. Single-Chain Soluble BG505.SOSIP gp140 Trimers as Structural and Antigenic Mimics of Mature Closed HIV-1 Env. *Journal of virology* 2015; 89(10):5318-5329.

151. Ködel Y. Virale Interferenz zwischen GBV-C und HIV: Einfluss N-terminaler Peptide des Glykoproteins E2 von GB-Virus C auf die HIV-Replikation. Südwestdeutscher Verlag für Hochschulschriften AG Company KG; 2011.
152. Yokoyama M, Nomaguchi M, Doi N, Kanda T, Adachi A, Sato H. In silico analysis of HIV-1 Env-gp120 reveals structural bases for viral adaptation in growth-restrictive cells. *Frontiers in Microbiology* 2016; 7.
153. Njai HF, Gali Y, Vanham G, Clybergh C, Jennes W, Vidal N, et al. The predominance of Human Immunodeficiency Virus type 1 (HIV-1) circulating recombinant form 02 (CRF02_AG) in West Central Africa may be related to its replicative fitness. *Retrovirology* 2006; 3(1):40.
154. Swetnam J, Shmelkov E, Zolla-Pazner S, Cardozo T. Comparative Magnitude of Cross-Strain Conservation of HIV Variable Loop Neutralization Epitopes. *PloS one* 2010; 5(12):e15994.
155. Hioe CE, Wrin T, Seaman MS, Yu X, Wood B, Self S, et al. Anti-V3 Monoclonal Antibodies Display Broad Neutralizing Activities against Multiple HIV-1 Subtypes. *PloS one* 2010; 5(4):e10254.
156. Almond D, Krachmarov C, Swetnam J, Zolla-Pazner S, Cardozo T. Resistance of Subtype C HIV-1 Strains to Anti-V3 Loop Antibodies. *Advances in Virology* 2012; 2012:8.
157. Pfeifer N, Walter H, Lengauer T. Association Between HIV-1 Coreceptor Usage and Resistance to Broadly Neutralizing Antibodies. *Journal of Acquired Immune Deficiency Syndromes (1999)* 2014; 67(2):107-112.
158. McCutchan FE. Global epidemiology of HIV. *Journal of medical virology* 2006; 78.
159. Santos RG, Appel JR, Giulianotti MA, Edwards BS, Sklar LA, Houghten RA, et al. The Mathematics of a Successful Deconvolution: A Quantitative Assessment of Mixture-Based Combinatorial Libraries Screened Against Two Formylpeptide Receptors. *Molecules (Basel, Switzerland)* 2013; 18(6):6408-6424.
160. Boggiano C, Jiang S, Lu H, Zhao Q, Liu S, Binley J, et al. Identification of a d-amino acid decapeptide HIV-1 entry inhibitor. *Biochemical and Biophysical Research Communications* 2006; 347(4):909-915.
161. Doranz BJ, Grovit-Ferbas K, Sharron MP, Mao S-H, Goetz MB, Daar ES, et al. A Small-molecule Inhibitor Directed against the Chemokine Receptor CXCR4 Prevents its Use as an HIV-1 Coreceptor. *The Journal of Experimental Medicine* 1997; 186(8):1395-1400.
162. Zhou N, Luo Z, Luo J, Fan X, Cayabyab M, Hiraoka M, et al. Exploring the Stereochemistry of CXCR4-Peptide Recognition and Inhibiting HIV-1 Entry with d-Peptides Derived from Chemokines. *Journal of Biological Chemistry* 2002; 277(20):17476-17485.
163. Eckert DM, Malashkevich VN, Hong LH, Carr PA, Kim PS. Inhibiting HIV-1 Entry: Discovery of D-Peptide Inhibitors that Target the gp41 Coiled-Coil Pocket. *Cell* 1999; 99(1):103-115.
164. Welch BD, Francis JN, Redman JS, Paul S, Weinstock MT, Reeves JD, et al. Design of a Potent d-Peptide HIV-1 Entry Inhibitor with a Strong Barrier to Resistance. *Journal of virology* 2010; 84(21):11235-11244.
165. Weinstock MT, Francis JN, Redman JS, Kay MS. Protease-Resistant Peptide Design – Empowering Nature's Fragile Warriors Against HIV. *Biopolymers* 2012; 98(5):431-442.
166. Chikaev AN, Bakulina AY, Burdick RC, Karpenko LI, Pathak VK, Ilyichev AA. Selection of peptide mimics of HIV-1 epitope recognized by neutralizing antibody VRC01. *PloS one* 2015; 10(3):e0120847.
167. Zwick MB, Bonnycastle LLC, Menendez A, Irving MB, Barbas CF, Parren PWHI, et al. Identification and Characterization of a Peptide That Specifically Binds the Human, Broadly Neutralizing Anti-Human Immunodeficiency Virus Type 1 Antibody b12. *Journal of virology* 2001; 75(14):6692-6699.
168. Saphire EO, Montero M, Menendez A, van Houten NE, Irving MB, Pantophlet R, et al. Structure of a High-affinity "Mimotope" Peptide Bound to HIV-1-neutralizing Antibody b12 Explains its Inability to Elicit gp120 Cross-reactive Antibodies. *Journal of molecular biology* 2007; 369(3):696-709.
169. Bock JE, Gavenonis J, Kritzer JA. Getting in shape: controlling peptide bioactivity and bioavailability using conformational constraints. *ACS chemical biology* 2013; 8(3):488-499.

Danksagung

Mein besonderer Dank gilt meiner Betreuerin Prof. Dr. Jutta Eichler für die Unterstützung und die Möglichkeit diese spannende Doktorarbeit in ihrer Arbeitsgruppe anfertigen zu dürfen. Die wissenschaftliche Arbeit auf dem Gebiet der Peptidforschung war für mich eine tolle horizonterweiternde Erfahrung und vielen Dank für die Motivation eigene Ideen einbringen zu können.

Ein herzlicher Dank geht an Prof. Dr. Barbara Schmitt für die Erstellung des Gutachtes und die spannende Kooperation mit der Möglichkeit die Peptide im HIV-1 Assay zu testen, die tolle Zusammenarbeit im Projekt und die neuen Einblicke in die Virologie. Bedanken möchte ich mich auch bei Tamara Rügamer, Johanna Schaubächer, Anette Rohrhofer für die Durchführung der HIV-1 Neutralisationsversuche und im besonderen Tamara für die Zusammenarbeit am Paper.

I want to thank Clemencia Pinilla and Jon Appel for the great opportunity to visit and work with them at Torrey Pines, for getting a good insight into the fascinating world of combinatorial libraries, for your welcoming and kind hospitality. I am very thankful that I had the chance to get know you and to work with you.

I want to thank Radleigh Santos for supporting me with the data analysis of the positional scanning results and for the good discussions on the results of the library scans.

Ein großer Dank gilt allen aktuellen und ehemaligen Kollegen meines Arbeitskreises mit denen ich zusammengearbeitet habe. Die gemeinsame Zeit war für mich eine tolle Erfahrung.

Danke an den Kaffeeraum und seine „Bewohner" ohne den ich meinen Reflux bestimmt nur halb so gut mit Kaffee trainiert hätte und die schöne Zeit.

Die Schuhstraße war für mich ein Ort, an dem ich sehr liebenswerte und mein Leben bereichernde Menschen kennengelernt habe von denen ich weiß, dass sie mich auch weiterhin begleiten werden. Ich danke besonders Christina und Jannis für ihre Unterstützung beim Schreiben, Marek für die positive Energie, für die Unterstützung nur eine Tür weiter danke Josi, Jasmin, Amelie und Dagmar. Und danke denen für ihren Support ohne den ich nicht den Weg nach und durch Erlangen gegangen wäre danke Peter, Erik und Robin. Für die vielen Stunden und die scheinbar nie endende Geduld - ich weiß es sehr zu schätzen, die ich in meinem Leben zu haben – Danke Dodo.