Original Article

Computational-guided determination of the functional role of 447-52D long CDRH3

Edwin Kamau¹, Richard Bonneau^{2,3}, and Xiang-Peng Kong^{1,*}

¹Department of Biochemistry and Molecular Pharmacology, NYU School of Medicine, New York NY 10016, USA,

²Department of Biology, Center for Genomics and Systems Biology and Computer Science Department, Courant Institute of Mathematical Sciences, New York University, New York NY 10003, USA, and ³Center for Computational Biology, Flatiron Institute, New York NY 10010, USA

*To whom correspondence should be addressed. E-mail: xiangpeng.kong@med.nyu.edu

Edited by: Valerie Daggett, Board Member for PEDS

Received 22 February 2019; Revised 6 March 2019; Editorial Decision 8 March 2019; Accepted 20 March 2019

Abstract

447-52D (447) is a human monoclonal antibody that recognizes a conserved epitope in the crown region of the third variable loop (V3) of HIV-1 gp120, and like many anti-HIV-1 antibodies with broad neutralization capabilities, it has a long heavy-chain complementarity determining region (CDRH3). Here, we use a combination of computational mutagenesis and modeling in tandem with fluorescence polarization assays to interrogate the molecular basis of 447 CDRH3 length and the individual contribution of selected CDRH3 residues to affinity. We observe that 447 CDRH3 length provides a large binding surface area and the best enthalpic contributions derived from hydrophobic packing, main-chain hydrogen bonds, electrostatic and van der Waals interactions. We also found out that CDRH3 residue Try^{100I} is critical to 447 binding affinity.

Key words: antibody, antigen-binding site, complementarity determining region

Introduction

HIV-1 entry into the host cell involves sequential interactions with the host cell-surface receptors CD4 and chemokine receptors (either CCR5 or CXCR4; Moore *et al.*, 1997; Berger *et al.*, 1999). Interactions with the host cell-surface receptors are mediated by two distinct regions located on the viral envelope surface glycoprotein, gp120. Consequently, the CD4 and chemokine binding sites on gp120 have conserved structural elements that facilitate binding to the host CD4 and chemokine receptors, respectively (Kwong *et al.*, 1998; Huang *et al.*, 2007; Shaik *et al.*, 2019). The chemokine receptor binding site is composed of the bridging sheet and the third variable region (V3) of gp120 and only accessible upon CD4 binding (Huang *et al.*, 2007; Shaik *et al.*, 2019). Despite its name, V3 is relatively conserved for its sequence and length in comparison with other variable regions on gp120 due to its indispensable role in viral entry and as a principal determinant of chemokine receptor usage (Moore *et al.*, 1997; Wyatt and Sodroski, 1998; Doms and Moore, 2000; Cardozo *et al.*, 2007). The V3 conserved structural elements necessitates viral fitness and are targeted by broadly

reactive antibodies (Cardozo *et al.*, 2007; Huang *et al.*, 2007; Jiang *et al.*, 2010). Nevertheless, such antibodies are ineffective in neutralizing field isolates of HIV-1 as the V3 is buried and only becomes accessible upon CD4 binding (Arendrup *et al.*, 1995). Therefore, the applicability of the V3 for vaccine development is rather limited.

Crystallographic structures of V3 in context of CD4-bound gp120 (Huang *et al.*, 2005; Shaik *et al.*, 2019), native gp120 trimer (Sanders *et al.*, 2013) and peptide in complex with anti-V3 antibodies (Stanfield and Wilson, 2005; Stanfield *et al.*, 2006; Jiang *et al.*, 2010) have shown that the V3 can be divided into three regions: the V3 base located proximal to the disulfide bond on the gp120 core, the V3 crown, located distal from V3 base and the V3 stem, a highly flexible region that connects the V3 base to the V3 crown (Huang *et al.*, 2005). The V3 crown region is highly immunogenic, eliciting antibodies during infection and after vaccination. Importantly, a number of broadly reactive anti-V3 monoclonal antibodies (mAbs) have been isolated and characterized in functional and structural studies (Conley *et al.*, 1994; Hioe *et al.*, 2010).

Structural analyses of V3 crown (gp120 residues 304–319, HXB2 numbering (Ratner *et al.*, 1987)) in complex with anti-V3 mAbs have revealed that although the V3 crown has conformational heterogeneity, it often forms a β -hairpin structure that can be divided into three subregions: the band, the circler and the arch (Jiang *et al.*, 2010). The three subregions display four conserved elements. The band has conserved positively charged residues 304 and 305 and a Tyr³¹⁸ in the N- and C-termini, respectively. The circler has a hydrophobic and hydrophilic face; the hydrophobic face contains two highly conserved isoleucines (Ile³⁰⁷ and Ile³⁰⁹) in the N-terminal strand and a conserved Phe³¹⁷ in the C-terminal strand. The hydrophilic face comprises of variable polar residues in positions 306, 308 and 316 in the N- and C-termini, respectively. Finally, the arch (residues 312–315) has a highly conserved four-residue sequence motif GPGX where the X residue is often an arginine (R) in clade B viruses and a glutamine (Q) in non-clade B viruses (Dhillon *et al.*, 2008; Burke *et al.*, 2009; Killikelly *et al.*, 2013). Structural studies have also revealed that anti-V3 mAbs recognize the conserved V3 crown elements in two distinct binding modes: the ‘ladle’ and ‘cradle’ modes (Jiang *et al.*, 2010). In the ladle mode, the shape of the antigen-binding site mimics a soup ladle with the arch of V3 pointing directly toward the antigen-binding site while in the cradle mode the V3 crown lies sideways in the long antigen-binding groove (Stanfield and Wilson, 2005; Stanfield *et al.*, 2006; Jiang *et al.*, 2010). Of the anti-V3 mAbs that target the V3 crown, the human mAb 447-52D (447) is the best studied and the focus of our investigation.

Structural studies demonstrated that 447 has a long chain complementarity determining region (CDRH3) (20 residues; Kabat definition (Kabat and Wu, 1991)) whose apex extends ~25 Å from the CDRH3 base. The C-terminal segment of CDRH3 apex is decorated with five consecutive tyrosines (Stanfield *et al.*, 2004; Killikelly *et al.*, 2013). The 447's antigen-binding site has a ladle-like shape that can be divided into two regions: the ‘handle’ and the ‘bowl’ regions (Burke *et al.*, 2009; Jiang *et al.*, 2010; Killikelly *et al.*, 2013). The ‘handle’ region, CDRH3, has a β -hairpin structure that extends upright and buries the V3 crown N-terminal half segment. An intermolecular parallel β -sheet interface is formed between the C-terminal strand of CDRH3 apex (residues H100F-100J) and the N-terminal strand of the V3 crown (residues 305–309). Interfacial side-chain pairing occurs on both sides of the β -sheet and has two distinct patches: (i) a hydrophobic patch consisting of CDRH3 residues Tyr^{H100G} and Tyr^{H100I} packing against the V3 circler hydrophobic face and V3 band residue 318 and (ii) a polar patch consisting of CDRH3 residues Asp^{H100F}, Tyr^{H100H} and Tyr^{H100J} interfacing with V3 crown residues 306 and 308. Additionally, the CDRH3 polar patch residues makes four main-chain hydrogen bonds with the conserved V3 crown hydrophobic face residues (Sharon *et al.*, 2003; Killikelly *et al.*, 2013). The ‘bowl’ region is a shallow pocket at the base of CDRH3 and consists of heavy-chain residues (Asp^{H95}, Tyr^{H100J} and Trp^{H33}) and light-chain residues (Trp^{L92} and Trp^{L96}) that sandwich the conserved V3 arch residues through side-chain interactions. The two light-chain residues pack against the GPG turn while the heavy-chain residues form a salt bridge and π -cation interactions with V3 crown residue Arg³¹⁵, respectively (Killikelly *et al.*, 2013).

Investigations by us and other groups have shown that the 447 epitope is located in the V3 arch and its specificity is determined by the highly specific binding of the Arg³¹⁵ side-chain. This explains the preferential neutralization of viruses harboring GPGR over those with GPGQ in their V3 arch (Stanfield *et al.*, 2004; Pantophlet

et al., 2007; Killikelly *et al.*, 2013). However, despite extensive functional and structural studies on 447 there remain some unanswered questions such as what is the necessity of the lengthy CDRH3? And which CDRH3 residues are critical to its binding affinity? Both questions can be addressed using truncation and alanine replacement. Furthermore, computational-guided mutational analysis can be performed prior to experimental evaluation to narrow down the experimental space to perturbations that maximize our understanding of this critical interaction.

Here, we leverage the idea that in most peptide–protein interactions the protein partner does not substantially change its structure upon binding as its interface is predefined and ready to accommodate the binding peptide (London *et al.*, 2010). We also assume that small changes to the 447 variable regions will not alter the overall (VL, VH) domain structure. Thus, we can determine the functional role of CDRH3 by performing *in silico* truncations using a computational protocol that only perturbs the CDRH3 region without altering the rest of 447 structure. To this end, we use RosettaRemodel to truncate and *de novo* loop model the truncated CDRH3 apex region (Huang *et al.*, 2011). This allows us to predict the effect of truncation on 447 stability and affinity to guide experimental investigation. We also perform computational alanine replacement mutations to identify and experimentally test CDRH3 apex residues potentially critical to 447 binding affinity. We found that 447 CDRH3 length is optimal to compensate for the loss of V3 peptide conformational entropy upon binding by providing a large binding surface area and great enthalpic contributions derived from hydrophobic packing, main-chain hydrogen bonds, van der Waals and electrostatic interactions. We also found out that CDRH3 apex residues Glu^{H100E} and Try^{H100I} are critical for binding.

Materials and methods

Computational truncation of CDRH3

We used the RosettaRemodel application on a fixed backbone (Huang *et al.*, 2011) to truncate the CDRH3 apex. The V3_{MN-CYCLIC} peptide in the crystal structure of the 447–V3_{MN-CYCLIC} complex (PDB ID: 4M1D) was retained in the same location in relation to the 447 variable region of heavy and light chains during the remodeling process. Briefly, the constant regions of the heavy and light-chain antigen-binding fragment (Fab) were first deleted from the crystal structure of the 447–V3_{MN-CYCLIC} complex. Then, cycles of packing and all-atom minimization were performed to energetically minimize any high energy regions in the starting crystal structure using the Rosetta relax protocol (Nivon *et al.*, 2013). During the minimization procedure, backbone and disulfide bond constraints were applied to reduce backbone structural deviation from the starting structure. An ensemble of 100 energetically minimized complex structures were generated, of which the lowest energy complex was then used to create truncated mutants via RosettaRemodel application (Huang *et al.*, 2011). The RosettaRemodel application requires, in addition to an energetically minimized structure, defined tasks that are to be performed during the remodeling procedure. A truncation task file was generated as described in Huang *et al.* (2011) so that 2-, 4-, 6- and 8-residue deletions in the CDRH3 region were independently generated. Secondary structures and residue assignments on the segment and the region flanking the deletion were included in the truncation task file to aid in the loop modeling process. The defined tasks directed the application to simultaneously perform: *in silico* deletion on the specified region, harvest fragments

from the PDB database with the assigned secondary structure to close the resulting break, design with the assigned residues and link the deleted segments via loop modeling using harvested fragments. Five lowest energy models were generated per mutant, of which the lowest energy model was selected for subsequent studies.

For each mutant (as well as the relaxed wildtype complex), an ensemble of 50 energetically minimized complexes was generated using the Rosetta relax application (Nivon *et al.*, 2013). With each complex, we computed the binding energy score, ΔG , using Rosetta Interface Analyzer application (Stranges and Kuhlman, 2013). $\Delta\Delta G$ was then calculated by taking the difference between ΔG of the truncated mutant from the average of the wt ΔG using the equation below:

$$\Delta\Delta G = \Delta G^{\text{mut}} - \text{mean}(\Delta G^{\text{wt}})$$

The interface analyzer application also computes the change in solvent accessible surface area (ΔSASA) of each interfacial residue using implementation of Lee and Richards NACCESS program (Hubbard and Thornton, 1993) (with a score broken into $\Delta\text{SASA}_{\text{Hydrophobic}}$ and $\Delta\text{SASA}_{\text{polar}}$ obtained by taking the difference between total ΔSASA and $\Delta\text{SASA}_{\text{Hydrophobic}}$). We also decomposed the rosetta all atom energy score terms (Alford *et al.*, 2017) in order to compute and compare interfacial residue–residue pair energies. The contribution of van der Waals (fa_{atr}), electrostatic (fa_{elec}), solvation (fa_{sol}) and hydrogen-bonding (hbond) energy terms were considered. Text describing the computational methods is in Supplementary information S1.

Computational-guided identification of 447 CDRH3 apex hot spot residues

In protein–protein interfaces, only a small number of residues make critical contribution to binding energy. These residues are termed as hot spots (Clackson and Wells, 1995). Hot spots are identified by mutating (computationally or experimentally) each interfacial residue to alanine and measuring the effect of mutation on binding affinity (Elcock *et al.*, 2001; Kortemme *et al.*, 2004; Moreira *et al.*, 2007). To identify 447 CDRH3 apex hot spots, we performed computational alanine scanning mutagenesis using the interface alanine scanning application in Robetta server (Kortemme *et al.*, 2004) to narrow down the number of residues to be investigated experimentally. The alanine scanning application automatically identifies the protein–protein interface, performs alanine replacement for each interfacial residue, computes the binding free energy (ΔG), and then computes the difference in binding energy ($\Delta\Delta G$) between the wild type and the alanine mutant to determine the contribution of the mutation to binding free energy of the protein–protein complex (Kortemme and Baker, 2002; Kortemme *et al.*, 2004). Residues (upon mutation to alanine) predicted to decrease $\Delta\Delta G$ by more than 1 kcal/mol are considered to be hot spots, and such CDRH3 apex residues were selected for experimental alanine scanning mutagenesis studies.

Site-directed mutagenesis

In order to experimentally generate truncated and individual alanine mutants, we used a pBR322 DNA plasmid encoding the 447 heavy-chain as a template for Site-directed mutagenesis using the QuickChange lightning mutagenesis kit (Agilent). For truncated mutants, primers were designed to delete two residues at a time. Each confirmed deleted mutant plasmid DNA was used as a template for the next deletion. Individual alanine or truncated mutations

were confirmed via sequencing (Macrogen, USA). Both the heavy and light-chain plasmid DNA were a gift from Dr Susan Zolla-Pazner, Icahn School of Medicine, Mt. Sinai University, New York, NY, USA.

Antibody expression

Each of the confirmed heavy-chain mutants or the wildtype DNA was co-transfected at a 1:1 heavy: light-chain ratio into FreeStyle 293-F cells (Human Embryonic Kidney cells; Invitrogen) using 25 kDa linear polyethylenimine (PEI; Polysciences Inc.) as the transfection reagent at a ratio of 3:1 PEI to DNA. About 500 ml culture was used for each mutant or wild type, and the supernatant was collected on Day 5. Each IgG was then purified from the supernatant on a packed protein G column using the AKTA purification system (GE healthcare). Each eluted IgG was dialyzed in PBS overnight, concentrated and stored either at -20 or -80°C until use. The expression level for all mutants was at a comparable level to that of the wild type as analyzed by SDS-PAGE. To produce Fabs of the IgGs, papain (Worthington, Lakewood, NJ) was activated with 10 mM cysteine and incubated with each purified IgG at weight ratio of 1:20 for 1 h at 37°C . The reaction was stopped by adding 10 mM iodoacetamide. Each Fab was then isolated from the fragment crystallizable (Fc) and undigested IgG by passing the reaction through a protein G column. Fab collected from the flowthrough was then purified by size-exclusion chromatography, dialyzed in PBS overnight, concentrated and stored in -80°C until use.

Enzyme-linked immunosorbent assays

Preliminary binding studies for the wild type and truncated mutants to V3_{MN-CYCLIC} (CRIHIGPGRAFYTC, Biomatik) (Killikelly *et al.*, 2013) and a monomeric gp120 of YU2 strain (a generous gift from Dr Xueling Wu, Aaron Diamond AIDS Research Center (ADARC), New York, NY) were determined by ELISA. Briefly, Immulon 4HBX plates were coated with 100 μl /well peptide at the concentration of 1.0 $\mu\text{g}/\text{ml}$ in phosphate buffered saline (PBS) and incubated overnight at 4°C . Plates were then washed three times with PBS and blocked with 200 μl /well of 1% bovine albumin serum (BSA) for 1 h at room temperature. Wells were then washed three times with PBS and incubated for 1 h with 3-fold diluted mutant or wildtype IgG (10 000–40 ng/ml) in PBS. The wells were washed again three times, and 100 μl /well of alkaline phosphate-conjugated goat anti-human IgG diluted 1:2000 in PBS was added and incubated for 1 h. After another three washes, 100 $\mu\text{l}/\text{ml}$ of diethalonamine buffer containing 10% *p*-nitrophenyl phosphate substrate was added, and the absorbance was read after 45 min at 405 nm.

Fluorescence polarization assays

Relative affinities of the wild type and mutants for V3 peptide were determined using the fluorescence polarization (FP) binding assay using a DTX 880 Multimode Detector (Beckman Coulter). Briefly, a 15-mer fluorescein isothiocyanate (FITC) conjugated V3 peptide from HIV-1 MN strain (MN-FITC) was synthesized by Biomatik to contain the amino acid sequence RRIHIGPGRAFYTTK-FITC. All samples were prepared in 96 black round bottom well plates in 0.1% pluronic acid F-68 (Sigma). 300 nM of the MN-FITC peptide probe was added to two-fold serially diluted Fab (6 μM –5 pM). The resulting solution was incubated at 25°C for 15–30 min with the excitation and emission wavelengths at 485 and 535 nm, respectively, before measuring the polarization. Binding constant (K_D)

values were determined using specific binding with the Hill slope model in Graphpad Prism version 7.0b, and the results were from averages of three independent experiments. To determine the statistical significance, one-way ANOVA with Bonferroni's multiple comparisons test was performed using GraphPad Prism version 7.0b for Mac.

Results

Truncated mutants of 447 CDRH3 maintain a β -hairpin structure and are predicted to have decreased affinities to V3

To investigate the contribution of the CDRH3 length to V3 binding affinity, we designed a truncation strategy that would sequentially delete one residue from each side of the CDRH3 β -hairpin (Fig. 1a) while retaining the β -turn residues Gly^{H100B}-Val^{H100C}-Ser^{H100D} (Fig. 1b). We first computationally modeled these truncations. Four

CDRH3 truncation mutants models, 447 Δ 2, 447 Δ 4, 447 Δ 6 and 447 Δ 8 representing 2-, 4-, 6- and 8-deletions, respectively, were generated using the RosettaRemodel application (Huang *et al.*, 2011). Visual inspection of the generated models showed that the CDRH3 for all the truncated mutants had a β -hairpin structure similar to the wild type (Fig. 1c–f). However, there were some notable differences from the wild type structure either in the number of hydrogen bonds or interfacial pairing between CDRH3 and V3. We found that 447 Δ 2 had four main-chain hydrogen bonds that were mediated by the same residues as the wild type (Fig. 1a and c). The 447 Δ 4 mutant also had four main-chain hydrogen bonds, however, Ser^{H100D} together with Tyr^{H100H}, and Tyr^{H100J} mediated the main-chain hydrogen bonds to V3 crown residues Ile³⁰⁷ and Ile³⁰⁹ (Fig. 1d). In contrast, both mutant 447 Δ 6 and 447 Δ 8 had only two hydrogen bonds to the V3 crown residue I309. One of the hydrogen bonds was mediated through Tyr^{H100J} in both mutants while the other was mediated through Tyr^{H100H} in 447 Δ 6 and Ser^{H100D} in 447 Δ 8 (Fig. 1e and f). Consequently, both of these mutants did not

Fig. 1 Truncation of CDRH3 of 447. (a) 447 complex structure showing the variable regions of the heavy (green) and light (cyan) chains and V3 crown (magenta) (PDB ID: 4M1D) (Killikelly *et al.*, 2013). Main-chain hydrogen bonds are shown and are presented as black dashed lines on the main figure and on inset figures. CDRH3 and V3 residues making hydrogen bonds are also indicated. Inset figures: The 447 'handle' region hydrophobic (top left) and polar (top right) patches. The 447 'bowl' region side-chain interactions (bottom). Interfacial residues in the 'bowl' or 'handle' region are shown as sticks, labeled by three-letter amino acid codes and residue numbers. (b) Sequence alignment of CDRH3 region showing the strategy used to shorten 447 CDRH3 apex. Gaps indicate where the deletions were made. β -turn residues Gly^{H100B}-Val^{H100C}-Ser^{H100D} were retained for each truncated mutant. Residue numbers are indicated above the alignment panel based on the Kabat numbering scheme (Kabat and Wu, 1991). (c–f) Structural models showing the backbone atoms of the CDRH3 region and V3 crown for the truncated mutants generated using RosettaRemodel (Huang *et al.*, 2011) 447 Δ 2 (c), 447 Δ 4 (d), 447 Δ 6 (e) and 447 Δ 8 (f). Main-chain hydrogen bonds for each mutant are shown and are presented as black dashed lines. CDRH3 and V3 residues making hydrogen bonds are also indicated. The truncated mutants maintain WT β -hairpin structure of 447 CDRH3.

maintain a β -sheet like interface due to the loss of two main-chain hydrogen bonds. Finally, we also inspected each mutant to determine the effect of truncations on interfacial pairing. The 447 Δ 2 and 447 Δ 4 mutants maintained the same side-chain pairing as in the wild type on the hydrophobic face, but only residues Tyr^{H100H} and Tyr^{H100J} on the CDRH3 polar face paired with V3 circlet hydrophilic face residues. We note that although CDRH3 polar face residue Asp^{H100F} was retained in the 447 Δ 2 mutant, it did not participate in pairing with V3 crown polar face residues because it was shifted from the CDRH3 C-terminal strand onto the β -turn. In contrast, both the 447 Δ 6 and 447 Δ 8 mutants maintained side-chain pairing between Tyr^{H100I} and V3 crown Ile³⁰⁹ only. Overall, although we observed several differences in interfacial pairing and main-chain hydrogen bonds, the truncation strategy predicted the generation of stable mutants with wild type β -hairpin structure.

Having established the structural effect of truncations on the CDRH3 apex and interfacial pairing, we then calculated and compared the change in binding energetics between the wild type and the truncated mutants. For each truncated mutant or energetically minimized wildtype complex, we generated an ensemble of 50 energetically minimized complexes. The generated complexes were then

used to compute the difference in binding energy ($\Delta\Delta G$) between the binding energy (ΔG) of the wild type and that of truncated mutant as described in the Materials and methods section. Commonly, $\Delta\Delta G > 1.0$ Rosetta energy unit (REU) predicts a mutation that decreases the binding energy, while $\Delta\Delta G < -1.0$ REU predicts a mutation that increases binding energy. $\Delta\Delta G$ score between -1 and 1 REU predicts a mutation with no effect on the binding energy. Using this criterion, all the truncated mutants had a decrease in binding energy, recording $\Delta\Delta G$ of 3, 2, 8 and 12 REU for 447 Δ 2, 447 Δ 4, 447 Δ 6 and 447 Δ 8 mutants, respectively (Fig. 2a).

We also calculated and compared the total solvent accessible surface buried in the interface (ΔS_{ASA}), and the interfacial change in electrostatic, van der Waals (ΔG_{atr}), solvation (ΔG_{sol}), electrostatic (ΔG_{elec}) and hydrogen bonding (ΔG_{hbond}) contributions to binding energy as described in the Materials and Methods section. ΔS_{ASA} decreased with increasing truncations with the wild type recording the largest buried surface area of 1273 $\text{\AA}^2$ while 447 Δ 2, 447 Δ 4, 447 Δ 6 and 447 Δ 8 buried 1107 $\text{\AA}^2$, 1089 $\text{\AA}^2$, 945 $\text{\AA}^2$ and 881 $\text{\AA}^2$, respectively (Fig. 2b). However, decomposition of the total buried surface areas into polar and hydrophobic components revealed a trend toward decrease in polar buried surface and an increase in

Fig. 2 Computational affinities of the WT and truncation mutants. (a) Computational binding energy ($\Delta\Delta G_{bind}$) predictions for the mutants and the wild type. Calculations were performed as described in material and methods, and the error bars were calculated from 50 models for the WT and each mutant. Computed total solvent accessible surface area buried in the interface. The total buried surface area has been decomposed into polar (grey) and hydrophobic components (black) (b), electrostatic (ΔG_{elec}) (c) and van der Waals (ΔG_{atr}) (d) contribution to binding energy at the 447-V3 interface. WT and the truncation mutants' computed calculations are shown. Direct correlation between decrease in binding energy and decrease in the contribution of all the computed energetics was observed that did not completely correlate with the increase in truncation.

Fig. 3 Expression and measured affinities of the WT and truncation mutants. (a) SDS-PAGE gel showing the expression of 447 WT and truncated mutants generated in this study. Top band heavy chain (~75 kDa), bottom band light chain (~25 kDa). (b) Binding affinities of the Fabs of the WT and truncated mutants to a linear V3_{MN-FITC} peptide determined by a fluorescence polarization assay. Truncated mutants have decreased binding affinities. Interestingly, 447Δ2 and 447Δ4 show similar binding affinities.

hydrophobic buried surface with increasing truncations. Thus, the wild type had the most polar buried surface and the least hydrophobic buried surface whereas the 447Δ2, 447Δ6 and 447Δ8 mutants had a decrease in polar buried surface and an increase in hydrophobic buried surface. However, the 447Δ4 mutant had a higher polar buried surface than all the other truncated mutants (Fig. 2b). The wild type also had the highest interfacial van der Waals energy contribution for the wild type (−37.1 kcal/mol) while the mutants recorded −25.9, −25.7, −18.8 and −14.8 kcal/mol for 447Δ2, 447Δ4, 447Δ6 and 447Δ8, respectively (Fig. 2c). The wild type also incurs the highest cost of removing interfacial contacting water molecules recording 18 kcal/mol while the truncated mutants had 5, 6, 8 and 10 less desolvation penalty for 447Δ2, 447Δ4, 447Δ6 and 447Δ8, respectively (Fig. 2d). The wild type also had the most interfacial electrostatic contribution of −5.1 kcal/mol while 447Δ2 (same for 447Δ4), 447Δ6 and 447Δ8 recorded 1.5, 2.7 and 3.4 kcal/mol, respectively, less electrostatic contribution than the wild type (Fig. 2e). As expected the hydrogen-bonding energy contribution for 447Δ2, 447Δ4 was similar to the wild type, slightly decrease for 447Δ6, and half as much for the 447Δ8 mutant (Fig. 2f). Thus, a direct correlation between length of truncation and decrease in binding energy, buried surface area, van der Waals, solvation and electrostatic energies was observed. While decrease in hydrogen-bonding energy was observed for 447Δ6 and 447Δ8 mutants.

Truncated mutants have decreased binding affinities

To experimentally measure the binding affinity of our truncation mutants, we generated each mutant by site-directed mutagenesis of the heavy chain of wild type 447, and co-expressed it with the wild type light chain in FreeStyle 293 F cells. Truncated mutants expressed at similar levels as the wild type, confirming that the truncations did not destabilize the mutant antibodies (Fig. 3a). With each mutant, we performed preliminary binding studies via ELISA using a cyclic V3 peptide, V3_{MN-CYCLIC} (CRIHIGPGRAFYTTC), to establish antigen recognition through binding activity. From this study, the 447Δ8 mutant had no binding activity (data not shown), so further studies with this mutant were not pursued. For the remaining truncated mutants, 447Δ2, 447Δ4 and 447Δ6, together with the wild type, we performed ELISA studies using monomeric

Table I. Fluorescence polarization data for truncated mutants

mAb	^a K _D (nM)	^b Fold change	^c P
447	9.3 ± 0.3	1.0	
447Δ2	210.3 ± 21.0	23.0	0.003
447Δ4	158.7 ± 9.0	17.0	0.0010
447Δ6	2377.0 ± 619	258.0	<0.0001

^aK_D values and error were obtained from three replicated experiments of mAb and linear V3_{MN-FITC} via fluorescence polarization using Graphpad Prism 7.0b specific binding with hill slope model.

^bChange in binding affinity calculated by taking the ratio of mutant over wild type K_D value.

^cAdjusted P-value at 95% confidence level comparing wild type over truncated mutant binding affinity levels. Values were obtained using one-way ANOVA with Bonferroni's multiple comparison test using Graphpad Prism 7.0b.

gp120 of YU2 strain (data not shown). In this context, the gp120 surface is glycosylated, allowing us to investigate the effect to binding affinity of the glycans neighboring 447–V3 interface. We found that the wild type and the remaining truncated mutants recognized the V3 crown with lower affinities than those recorded with cyclic V3 peptide. Having established the mutants also bound the V3 crown in gp120 context, we embarked on establishing the strength of binding for the remaining mutants and wild type using FP assay. For this assay, a fluorescein-labeled linear V3 peptide, V3_{MN-FITC} (RRIHIGPGRAFYTTC-FITC) was used. We found that wild type 447 had a binding affinity (K_D) of 9.3 nM, close to the previously observed binding affinity of 23 nM using isothermal titration calorimetry (Killikelly *et al.*, 2013). Truncated mutants demonstrated lower binding affinities compared with the wild type, with 447Δ2 recording 210 nM while 447Δ4 and 447Δ6 recorded 158.7 nM and 2377 nM, respectively (Fig. 3b and Table 1). This resulted in 23-, 17- and 258-fold decreases in affinity, respectively, compared with the wild type (Table 1 and Fig. 3b), and showed a similar trend as observed in the computational binding studies (Fig. 2a). Taken together, truncating the CDRH3 apex led to generation of stably expressed mutants that bound (with the exception of the 447Δ8 mutant) V3 peptides and monomeric gp120. However, the truncated

mutants, like the wild type, bind monomeric gp120 with lower affinity than they do the cyclic V3 peptide perhaps due to steric clashes with neighboring glycans as previously observed for the wild type (Upadhyay *et al.*, 2014; Kumar *et al.*, 2015). The truncated mutants also demonstrate weaker binding strength than the wild type due to the changes in interfacial pairing along with decreases in buried surface area, electrostatic and van der Waal interactions.

Table II. Alanine scanning prediction and fluorescence polarization data

^a mAb	^b $\Delta\Delta G_{\text{bind}}$ (kcal/mol)	^c K_D (nM)	^d Fold change	^e P
447		9.3 ± 0.3	1.0	
447-Y ^{H98} A	1.10	5.3 ± 0.5	0.6	0.99
447-Y ^{H100G} A	1.23	14.4 ± 0.8	1.5	0.99
447-Y ^{H100H} A	1.39	9.6 ± 0.8	1.0	0.99
447-Y ^{H100I} A	2.54	61.2 ± 4.5	6.5	0.012

^aIndividual mutants named using one letter amino acid code; Isoleucine (I), Tyrosine (Y), Alanine (A).

^bAlanine scanning mutagenesis prediction obtained using Rosetta interface alanine scanning application.

^c K_D values and error were obtained from three replicated experiments of mAb and linear V3_{MN-FITC} via fluorescence polarization using Graphpad Prism 7.0b specific binding with hill slope model.

^dChange in binding affinity calculated by taking the ratio of mutant over wild type K_D value

^eAdjusted *P*-value at 95% confidence level comparing wild type over truncated mutant binding affinity levels. Values were obtained using one-way ANOVA with Bonferroni's multiple comparison test using Graphpad Prism 7.0b.

Residue Tyr^{H100I} contributes significantly to the binding affinity

Having established the relative importance of CDRH3 length to binding affinity, we then evaluated the relative importance of CDRH3 interfacial residues interacting with the V3 crown N-terminal half. To do this, we performed computational alanine scanning application in Rosetta (Kortemme *et al.*, 2004) to narrow down on which CDRH3 apex residues to focus on for experimental investigation. The alanine scanning application suggested seven CDRH3 residues, Asp^{H95}, Ile^{H98}, Tyr^{H100G}, Tyr^{H100H}, Tyr^{H100I}, Tyr^{H100J} and Tyr^{H100K}, to have $\Delta\Delta G$ greater than 1 kcal/mol. We selected the CDRH3 residues Ile^{H98}, Tyr^{H100G}, Tyr^{H100H} and Tyr^{H100I} for site-directed alanine scanning experiments and did not pursue the other predicted residues as we have previously characterized their individual contribution to binding affinity and neutralization (Killikelly *et al.*, 2013). We found that the Ile^{H98}Ala mutant resulted in a modest increase in affinity, while the Tyr^{H100G}Ala mutant had an opposite effect, showing a modest 1.5-fold decrease in affinity. Meanwhile, the Tyr^{H100H}Ala mutant displayed wild type affinity levels (Table 2, Fig. 4a–c). The results for these three mutants contrasted with the computational alanine scanning predictions and suggest that Rosetta overestimated their contribution to binding energy. Rosetta perhaps gave too much consideration to their hydrophobic packing and ignored the main-chain hydrogen-bonding contribution to binding. However, the prediction for CDRH3 apex residue Tyr^{H100I}Ala mutant correlated with experimental testing, as the mutant displayed a significant six-fold decrease in affinity (Table 2, Fig. 4d). These results suggest that the CDRH3 residue Tyr^{H100I} side-chain is a hot spot and likely anchor the CDRH3 apex onto the V3 crown N-terminus by packing against

Fig. 4 Effect of individual CDRH3 apex alanine mutants. (a–d) Binding affinity data for individual alanine mutants of key residues, Ile^{H98} (a) Tyr^{H100H} (b), Tyr^{H100G} (c) and Tyr^{H100I} (d), to the linear V3_{MN-FITC} peptide determined by the fluorescence polarization assay. Residues predicted to be hot spots via Rosetta alanine scanning mutagenesis application. Error bars were computed from three independent measurements. Note that only the Tyr^{H100I}Ala mutation had a measurable affinity difference.

the V3 residue Ile³⁰⁹. Each of the other three residues Ile^{H98}, Tyr^{H100G} and Tyr^{H100H} seem to only make minimal contribution to binding, but do suggest that some of the CDRH3 apex positions either do not have the optimal residue for binding affinity or are by design able to tolerate changes in side-chain size without significantly compromising binding affinity.

Discussion

Extralong CDRH3s can play important roles in antigen recognition for human antibodies; they are often found in mAbs isolated in HIV-1 patients and these anti-HIV-1 Env antibodies provide an unique opportunity for structural understanding of the functional roles of long CDRH3s. By a combination of computational prediction and experimental verification, we have systematically investigated the structure-function of the long CDRH3 of mAb 447, an archetype anti-HIV-1 V3 mAb. We found that its long CDRH3 provides a large binding surface optimal to stabilize the highly flexible V3 crown and two CDRH3 residues, Glu^{H100E} located at the tip of CDRH3 loop and Tyr^{H100I} residing on the side of CDRH3, that are critical for 447 binding affinity.

Long CDRH3s have been found to play distinct roles in different anti-HV1 Env antibodies. Many of them harbor an extended long CDRH3, standing in the mid of the antigen-binding site like that in 447. Most of them use the extended CDRH3 to reach between glycans covering the Env surface to touch the amino acids on the surface of the Env protein. For example, two clonally related broadly neutralizing mAb PG9 and PG16 that recognize glycan epitopes in the V1V2 region of HIV-1 Env each has a long protruding CDRH3 shaped like a hammerhead (Pejchal et al., 2010). These bnAbs use the protruding CDRH3 to reach between glycans located in the V1V2 region of the Env trimer so that the β -strand of the hammerhead can interact with a V2 β -strand by a β -sheet like interaction (McLellan et al., 2011; Pancera et al., 2013). Another family of so-called trimer specific bnAbs, including PGT145 and PGDM1400, have a straight CDRH3 as a long β -hairpin. They use this long protruding CDRH3 to reach between the apex glycans and into a hole formed at the very center of the Env trimer (Sok et al., 2014; Lee et al., 2017; Liu et al., 2017). Unlike these bnAbs, 447 is in a unique class that uses the long CDRH3 to interact with the extended variable loop using the side of its CDRH3 hairpin. Because the extended long strand-strand interaction between 447 CDRH3 and the V3 crown, it is possible to shorten slightly the length of its CDRH3 and it still maintains a substantial binding affinity. While in the cases of those bnAbs like PG9 and PGT145, the length of their CDRH3 is critical for their function. The 447 has a well-constructed antigen-binding site ideally designed to interact with the V3 crown with the ladle binding mode (Jiang et al., 2010; Killikelly et al., 2013). One may speculate that 447 type Abs happens rarely. However, our unpublished data have shown that mAbs like 447 can be isolated from different individuals, thus this type of mAbs may not be rare.

The 447 also provides a unique example for how tyrosine plays key roles in antigen-binding sites. Tyrosine is the most abundant amino acid at antigen-binding sites, accounting for 40% in CDRH3 of naive antibodies and can make more than 25% of the antigen contacts (Koide and Sidhu, 2009; Birtalan et al., 2010); this may stem from its large size and flat surface, ability to form hydrophobic, polar, and electrostatic (π - π or π -cation) interactions, and also an inflexible side-chain. These properties allow the tyrosine residue to accommodate diverse antigen surfaces ranging from hydrophobic, polar or charged, with high affinity and specificity (Koide and

Sidhu, 2009). The 447 harbors five tyrosines consecutively lining up, from top to bottom, one of the two CDRH3 strands, and the side chains of four of them (Tyr^{H100G}-Tyr^{H100J}) have direct contacts with the V3 epitope, while the last one (Tyr^{H100K}) is located on the side facing away from the epitope. This consecutively alignment of tyrosines along one strand creates a large surface area for the epitope to pack against. Tyr^{H100J}, together with Trp^{L91}, sandwiches the side chain of the V3 arch residue Arg³¹⁵ and its function has been carefully delineated previously (Killikelly et al., 2013). Interestingly, Tyr^{H100I}, which has a lower epitope surface contact area than that of Tyr^{H100H}, plays a key role in its binding affinity. A possible explanation of the observed decrease in affinity upon the alanine mutation is that side chain of Tyr^{H100I}, together with that of Tyr^{L32}, forms a boxed hydrophobic environment to pack the conserved V3 residue Ile³⁰⁹ and removing the side chain of Tyr^{H100I} will open this box and substantially decrease 447's affinity to V3.

Due to the almost unlimited possibilities for forming CDRs, it is still extremely hard to computationally predict the structure-function of the antigen-binding site of an antibody. Our method, as presented here and published previously (Killikelly et al., 2013), in dissecting the individual structural elements in the antigen-binding site of the canonical HIV-1 V3 mAb 447 by systematically interrogating amino acids in each element serves as an example of understanding an antigen-binding site.

Supplementary Data

Supplementary data are available at *Protein Engineering, Design and Selection* online.

Acknowledgements

We thank Dr Paramjit Arora and his lab members for helping with the fluorescence polarization assay, Dr Xueling Wu at ADARC for providing us with gp120 of YU2 strain, and Christina Luo for carefully reading the manuscript. This study was supported in part by NIH grant AI100151.

References

- Alford, R.F., Leaver-Fay, A., Jeliazkov, J.R., O'Meara, M.J., DiMaio, F.P., Park, H., Shapovalov, M.V., Renfrew, P.D., Mulligan, V.K., Kappel, K. et al. (2017) *J. Chem. Theory Comput.*, **13**, 3031–3048. doi:10.1021/acs.jctc.7b00125.
- Arendrup, M., Akerblom, L., Heegaard, P.M., Nielsen, J.O. and Hansen, J.E. (1995) *Arch. Virol.*, **140**, 655–670.
- Berger, E.A., Murphy, P.M. and Farber, J.M. (1999) *Annu. Rev. Immunol.*, **17**, 657–700. doi:10.1146/annurev.immunol.17.1.657.
- Birtalan, S., Fisher, R.D. and Sidhu, S.S. (2010) *Mol. Biosyst.*, **6**, 1186–1194. doi:10.1039/b927393j.
- Burke, V., Williams, C., Sukumaran, M., Kim, S.S., Li, H., Wang, X.H., Gorny, M.K., Zolla-Pazner, S. and Kong, X.P. (2009) *Structure*, **17**, 1538–1546. doi:10.1016/j.str.2009.09.012.
- Cardozo, T., Kimura, T., Philpott, S., Weiser, B., Burger, H. and Zolla-Pazner, S. (2007) *AIDS Res. Hum. Retroviruses*, **23**, 415–426. doi:10.1089/aid.2006.0130.
- Clackson, T. and Wells, J.A. (1995) *Science*, **267**, 383–386.
- Conley, A.J., Gorny, M.K., Kessler, J.A., 2nd, Boots, L.J., Ossorio-Castro, M., Koenig, S., Lineberger, D.W., Emini, E.A., Williams, C. and Zolla-Pazner, S. (1994) *J. Virol.*, **68**, 6994–7000.
- Dhillon, A.K., Stanfield, R.L., Gorny, M.K., Williams, C., Zolla-Pazner, S. and Wilson, I.A. (2008) *Acta Crystallogr. D Biol. Crystallogr.*, **D64**, 792–802. doi:10.1107/S0907444908013978.
- Doms, R.W. and Moore, J.P. (2000) *J. Cell Biol.*, **151**, F9–F14.

- Elcock,A.H., Sept,D. and McCammon,J.A. (2001) *J. Phys. Chem. B*, **105**, 1504–1518. doi: 10.1021/jp003602d.
- Hioe,C.E., Wrin,T., Seaman,M.S., Yu,X., Wood,B., Self,S., Williams,C., Gorny,M.K. and Zolla-Pazner,S. (2010) *PLoS One*, **5**, e10254. doi:10.1371/journal.pone.0010254.
- Huang,P.S., Ban,Y.E., Richter,F., Andre,I., Vernon,R., Schief,W.R. and Baker,D. (2011) *PLoS One*, **6**, e24109. doi:10.1371/journal.pone.0024109.
- Huang,C.C., Lam,S.N., Acharya,P., Tang,M., Xiang,S.H., Hussan,S.S., Stanfield,R.L., Robinson,J., Sodroski,J., Wilson,I.A. *et al.* (2007) *Science*, **317**, 1930–1934. doi:10.1126/science.1145373.
- Huang,C.C., Tang,M., Zhang,M.Y., Majeed,S., Montabana,E., Stanfield,R.L., Dimitrov,D.S., Korber,B., Sodroski,J., Wilson,I.A. *et al.* (2005) *Science*, **310**, 1025–1028. doi:10.1126/science.1118398.
- Hubbard,S.J. and Thornton,J.M. (1993) ‘NACCESS’, computer program. First published on, doi: citeulike-article-id:3431829.
- Jiang,X., Burke,V., Totrov,M., Williams,C., Cardozo,T., Gorny,M.K., Zolla-Pazner,S. and Kong,X.P. (2010) *Nat. Struct. Mol. Biol.*, **17**, 955–961. doi:10.1038/nsmb.1861.
- Kabat,E.A. and Wu,T.T. (1991) *J. Immunol.*, **147**, 1709–1719.
- Killikelly,A., Zhang,H.T., Spurrier,B., Williams,C., Gorny,M.K., Zolla-Pazner,S. and Kong,X.P. (2013) *Biochemistry*, **52**, 6249–6257. doi:10.1021/bi400645e.
- Koide,S. and Sidhu,S.S. (2009) *ACS Chem. Biol.*, **4**, 325–334. doi:10.1021/cb800314v.
- Kortemme,T. and Baker,D. (2002) *Proc.Natl. Acad. Sci. U.S.A.*, **99**, 14116–14121. doi:10.1073/pnas.202485799.
- Kortemme,T., Kim,D.E. and Baker,D. (2004) *Sci. STKE*, **2004**, pl2. doi:10.1126/stke.2192004pl2.
- Kumar,R., Pan,R., Upadhyay,C., Mayr,L., Cohen,S., Wang,X.H., Balasubramanian,P., Nadas,A., Seaman,M.S., Zolla-Pazner,S. *et al.* (2015) *J. Virol.*, **89**, 9090–9102. doi:10.1128/JVI.01280-15.
- Kwong,P.D., Wyatt,R., Robinson,J., Sweet,R.W., Sodroski,J. and Hendrickson,W.A. (1998) *Nature*, **393**, 648–659. doi:10.1038/31405.
- Lee,J.H., Andrabi,R., Su,C.Y., Yasmeen,A., Julien,J.P., Kong,L., Wu,N.C., McBride,R., Sok,D., Pauthner,M. *et al.* (2017) *Immunity*, **46**, 690–702. doi:10.1016/j.immuni.2017.03.017.
- Liu,Q., Acharya,P., Dolan,M.A., Zhang,P., Guzzo,C., Lu,J., Kwon,A., Gururani,D., Miao,H., Bylund,T. *et al.* (2017) *Nat. Struct. Mol. Biol.*, **24**, 370–378. doi:10.1038/nsmb.3382.
- London,N., Movshovitz-Attias,D. and Schueler-Furman,O. (2010) *Structure*, **18**, 188–199. doi:10.1016/j.str.2009.11.012.
- McLellan,J.S., Pancera,M., Carrico,C., Gorman,J., Julien,J.P., Khayat,R., Louder,R., Pejchal,R., Sastry,M., Dai,K. *et al.* (2011) *Nature*, **480**, 336–343. doi:10.1038/nature10696.
- Moore,J.P., Trkola,A. and Dragic,T. (1997) *Curr. Opin. Immunol.*, **9**, 551–562.
- Moreira,I.S., Fernandes,P.A. and Ramos,M.J. (2007) *J. Comput. Chem.*, **28**, 644–654. doi:10.1002/jcc.20566.
- Nivon,L.G., Moretti,R. and Baker,D. (2013) *PLoS One*, **8**, e59004. doi:10.1371/journal.pone.0059004.
- Pancera,M., Shahzad-Ul-Hussan,S., Doria-Rose,N.A., McLellan,J.S., Bailer,R.T., Dai,K., Loesgen,S., Louder,M.K., Staupe,R.P., Yang,Y. *et al.* (2013) *Nat. Struct. Mol. Biol.*, **20**, 804–813. doi:10.1038/nsmb.2600.
- Pantophlet,R., Aguilar-Sino,R.O., Wrin,T., Cavacini,L.A. and Burton,D.R. (2007) *Virology*, **364**, 441–453. doi:10.1016/j.virol.2007.03.007.
- Pejchal,R., Walker,L.M., Stanfield,R.L., Phogat,S.K., Koff,W.C., Poignard,P., Burton,D.R. and Wilson,I.A. (2010) *Proc. Natl. Acad. Sci. USA*, **107**, 11483–11488. doi:10.1073/pnas.1004600107. doi:10.1073/pnas.1004600107 [pii].
- Ratner,L., Fisher,A., Jagodzinski,L.L., Mitsuya,H., Liou,R.S., Gallo,R.C. and Wong-Staal,F. (1987) *AIDS Res. Hum. Retroviruses*, **3**, 57–69. doi:10.1089/aid.1987.3.57.
- Sanders,R.W., Derking,R., Cupo,A., Julien,J.P., Yasmeen,A., de Val,N., Kim,H.J., Blattner,C., de la Pena,A.T., Korzun,J. *et al.* (2013) *PLoS Pathog.*, **9**, e1003618. doi:10.1371/journal.ppat.1003618.
- Shaik,M.M., Peng,H., Lu,J., Rits-Volloch,S., Xu,C., Liao,M. and Chen,B. (2019) *Nature*, **565**, 318–323. doi:10.1038/s41586-018-0804-9.
- Sharon,M., Kessler,N., Levy,R., Zolla-Pazner,S., Gorlach,M. and Anglister,J. (2003) *Structure*, **11**, 225–236.
- Sok,D., van Gils,M.J., Pauthner,M., Julien,J.P., Saye-Francisco,K.L., Hsueh,J., Briney,B., Lee,J.H., Le,K.M., Lee,P.S. *et al.* (2014) *Proc.Natl. Acad. Sci. U.S.A.*, **111**, 17624–17629. doi:10.1073/pnas.1415789111.
- Stanfield,R.L., Gorny,M.K., Williams,C., Zolla-Pazner,S. and Wilson,I.A. (2004) *Structure*, **12**, 193–204. doi:10.1016/j.str.2004.01.003.
- Stanfield,R.L., Gorny,M.K., Zolla-Pazner,S. and Wilson,I.A. (2006) *J. Virol.*, **80**, 6093–6105. doi:10.1128/JVI.00205-06.
- Stanfield,R.L. and Wilson,I.A. (2005) *Hum. Antibodies*, **14**, 73–80.
- Stranges,P.B. and Kuhlman,B. (2013) *Protein Science*, **22**, 74–82. doi:10.1002/pro.2187.
- Upadhyay,C., Mayr,L.M., Zhang,J., Kumar,R., Gorny,M.K., Nadas,A., Zolla-Pazner,S. and Hioe,C.E. (2014) *J. Virol.*, **88**, 12853–12865. doi:10.1128/JVI.02125-14.
- Wyatt,R. and Sodroski,J. (1998) *Science*, **280**, 1884–1888.