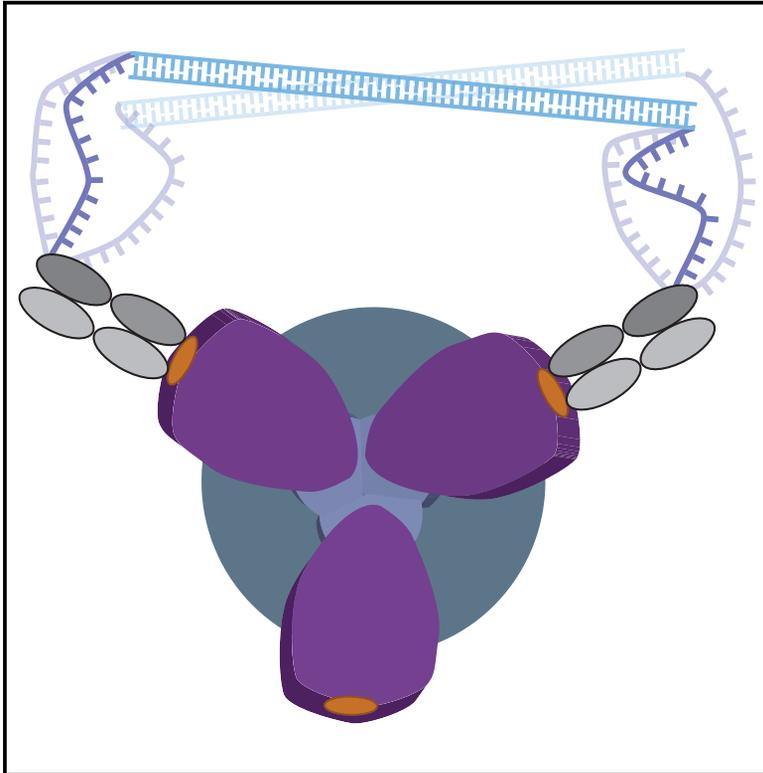


# Cell Systems

## Harnessing Avidity: Quantifying the Entropic and Energetic Effects of Linker Length and Rigidity for Multivalent Binding of Antibodies to HIV-1

### Graphical Abstract



### Highlights

- Synthetic antibodies that bivalently bind to HIV-1 can markedly enhance avidity
- Linkers that enable bivalent binding are fully characterized by the linker entropy
- Properly sized rigid linkers outperform long, flexible linkers
- Avidity can be further enhanced by increasing antibody valency

### Authors

Tal Einav, Shahrzad Yazdi, Aaron Coey, Pamela J. Bjorkman, Rob Phillips

### Correspondence

[bjorkman@caltech.edu](mailto:bjorkman@caltech.edu) (P.J.B.),  
[phillips@pboc.caltech.edu](mailto:phillips@pboc.caltech.edu) (R.P.)

### In Brief

Antibodies enhance their functional affinity by binding an antigen at multiple sites. Because HIV-1 contains few binding targets, the ability of antibodies to neutralize the virus is greatly decreased. Here, we develop a statistical-mechanical model that predicts the efficacy of synthetic antibodies composed of two Fabs joined together by DNA linkers of different lengths and flexibilities. This model predicts a synthetic antibody's neutralization potency based on its linker composition, providing a framework to guide the design of future multivalent therapies.

# Harnessing Avidity: Quantifying the Entropic and Energetic Effects of Linker Length and Rigidity for Multivalent Binding of Antibodies to HIV-1

Tal Einav,<sup>1</sup> Shahrzad Yazdi,<sup>2</sup> Aaron Coey,<sup>3</sup> Pamela J. Bjorkman,<sup>3,\*</sup> and Rob Phillips<sup>1,3,4,5,\*</sup>

<sup>1</sup>Department of Physics, California Institute of Technology, Pasadena, CA 91125, USA

<sup>2</sup>Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>3</sup>Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA

<sup>4</sup>Department of Applied Physics, California Institute of Technology, Pasadena, CA 91125, USA

<sup>5</sup>Lead Contact

\*Correspondence: [bjorkman@caltech.edu](mailto:bjorkman@caltech.edu) (P.J.B.), [phillips@pboc.caltech.edu](mailto:phillips@pboc.caltech.edu) (R.P.)

<https://doi.org/10.1016/j.cels.2019.09.007>

## SUMMARY

IgG antibodies increase their apparent affinities by using both of their Fabs to simultaneously attach to antigens. HIV-1 foils this strategy by having few, and highly separated, Envelope (Env) spike targets for antibodies, forcing most IgGs to bind monovalently. Here, we develop a statistical mechanics model of synthetic diFabs joined by DNA linkers of different lengths and flexibilities. This framework enables us to translate the energetic and entropic effects of the linker into the neutralization potency of a diFab. We demonstrate that the strongest neutralization potencies are predicted to require a rigid linker that optimally spans the distance between two Fab binding sites on an Env trimer and that avidity can be further boosted by incorporating more Fabs into these constructs. These results inform the design of multivalent anti-HIV-1 therapeutics that utilize avidity effects to remain potent against HIV-1 in the face of the rapid mutation of Env spikes.

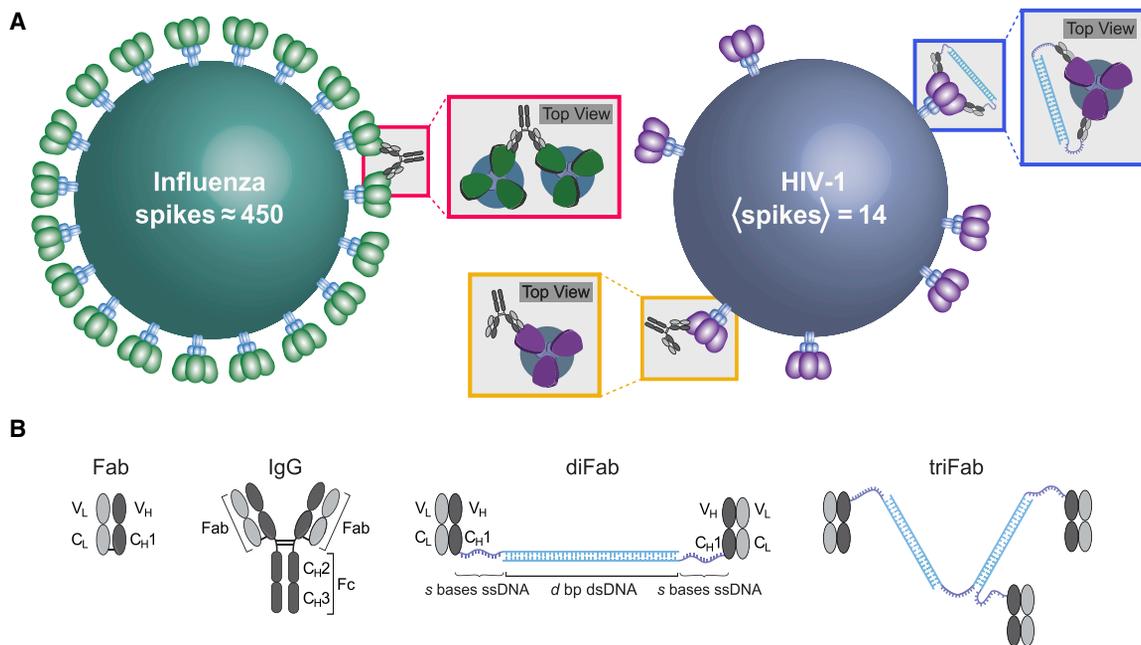
## INTRODUCTION

Despite decades of research since its discovery, human immunodeficiency virus-1 (HIV-1) continues to threaten global public health (UNAIDS, 2019). While there have been advances in our understanding of the mechanisms of infection and the development of preventative and therapeutic strategies, there remains no cure for HIV-1 infection. Antiretroviral therapy with small molecule drugs can control the progression of the virus, allowing those infected with HIV-1 to live longer and healthier lives, but the treatment includes detrimental side effects, and when discontinued or not taken as prescribed, leads to viral rebound to pre-treatment levels (Fauci, 2008). A major factor confounding the development of a prophylactic vaccine is the rapid mutation of HIV-1, which leads to the emergence of many new strains, even within a single individual (Escolano et al., 2017). Thus, most antibodies raised by the host immune system are strain-

specific or neutralize only a subset of strains, leading to viral escape from host antibodies.

Recent interest has focused upon the isolation of broadly neutralizing immunoglobulin G (IgG) antibodies (bNAbs) from a subset of HIV-1-infected individuals (McCoy and Burton, 2017). These antibodies bind to and block the functions of the HIV-1 Envelope (Env) spike, the viral protein responsible for the fusion of HIV-1 to the host cell (Klasse, 2012). The discovery and characterization of HIV-1 bNAbs has brought new impetus to the idea of passively delivering antibodies to protect against or treat HIV-1 infection. bNAbs can prevent and treat infection in animal models (Baba et al., 2000; Mascola et al., 2000; Hessel et al., 2007, 2009; Klein et al., 2012, 2013, 2014) and exhibited efficacy against HIV-1 in human trials (Caskey et al., 2015, 2017; Lynch et al., 2015; Schoofs et al., 2016). However, HIV-1 Env mutates to become resistant to any single bNAb, as even the most potent NAbs developed in an infected individual normally fail to neutralize autologous circulating viral strains (Moore et al., 2012; Liao et al., 2013; Wibmer et al., 2013; Doria-Rose et al., 2014). As a result, antibodies that develop during HIV-1 infection appear to be unable to control the virus in an infected individual.

We previously proposed that one mechanism by which HIV-1 evades antibodies more successfully than other viruses arises from the low surface density of Env spikes that can be targeted by neutralizing antibodies (Klein and Bjorkman, 2010; Galimidi et al., 2015). Compared to viruses such as influenza A, dengue, and hepatitis B, the density of Env spikes on the surface of HIV-1 is about 2 orders of magnitude smaller (Klein and Bjorkman, 2010). For example, influenza A has  $\approx 450$  spikes per virion, whereas each HIV-1 virion incorporates only 7–30 Env spikes (average of 14) (Chertova et al., 2002; Zhu et al., 2003, 2006; Liu et al., 2008; Klein and Bjorkman, 2010), even though both influenza A and HIV-1 are enveloped viruses with  $\approx 120$ -nm diameters (Figure 1A). The HIV-1 spikes are the machinery by which the virus binds its host receptor CD4 and coreceptor CCR5/CXCR4 to mediate the fusion of the host and viral membranes that allows its genome to enter target cells (Klasse, 2012). As a consequence of its small number of spikes, HIV-1 infection of target cells is inefficient; the transmission probabilities for sexually acquired HIV-1 infection range from 0.4% to 1.4% (Patel et al., 2014). However, the reduced infectivity of



**Figure 1. Effects of Spike Density on IgG Binding**

Figure360► For a Figure360 author presentation of this figure, see <https://doi.org/10.1016/j.cels.2019.09.007>. (A) Close spacing of surface spikes on influenza A allows bivalent binding of IgGs to adjacent spikes (red boxes). In contrast, HIV-1 has few spikes (14 on average) spaced far apart, and because the HIV-1 spike architecture prohibits simultaneous binding of two Fabs to a single Env trimer, most IgGs bind monovalently to HIV-1 Envs (gold boxes). We investigated a synthetic diFab designed to bind bivalently to a single HIV-1 spike trimer (blue boxes). (B) Schematics of a Fab, an IgG, a diFab composed of two Fabs joined together by  $d$  bp dsDNA, and two segments of ssDNA bases, and a triFab made up of three Fabs.

HIV-1 comes with a concomitant reduction in the ability of antibodies to control the virus, as the surface spikes serve as the only targets for neutralizing antibodies that can block infection of target cells (McCoy and Burton, 2017).

While the close spacing of spikes on typical viruses allows IgG antibodies to bind bivalently to neighboring spikes (inter-spike crosslinking) using both of their antigen-binding arms (Fabs), most HIV-1 spikes are too far apart (typically over 20 nm separation) (Klein and Bjorkman, 2010) to permit inter-spike crosslinking by IgGs whose antigen-binding sites are separated by  $\leq 15$  nm (Saphire et al., 2001). Furthermore, although each homotrimeric HIV-1 spike includes three binding sites (epitopes) for an antibody, the architecture of HIV-1 Envs prohibits simultaneous binding of two Fabs within a single IgG to the same Env (intra-spike crosslinking) (Klein, 2009; Wang et al., 2017). We suggested that predominantly monovalent binding by anti-HIV-1 antibodies expands the range of Env mutations permitting antibody evasion, since reagents capable of bivalent binding through inter- or intra-spike crosslinking would be less affected by Env mutations that reduce but do not abrogate binding and thus may be more potent across multiple strains of HIV-1 (Klein and Bjorkman, 2010; Galimidi et al., 2015). The hypothesis that HIV's low spike numbers and low densities contributes to the vulnerability of HIV-1 bNAbs to spike mutations is supported by independent biochemical and EM studies demonstrating that HIV-1 has an unusually low number of spikes that are not clustered (Layne et al., 1992; Chertova et al., 2002; Zhu et al., 2003, 2006; Liu et al.,

2008), and that bivalent IgG forms of anti-HIV-1 NAbs are only modestly more effective than monovalent Fabs, by contrast to bivalent IgGs against other viruses, which can be 100s- to 1,000s-fold more potent than counterpart monovalent Fabs (Klein, 2009; Klein and Bjorkman, 2010; Galimidi et al., 2015; Wang et al., 2017).

An antibody's neutralization potency against a virus is related to its antigen-binding affinity, which is defined as the binding strength between a Fab and its antigen (Eisen and Siskind, 1964) described by the equilibrium dissociation constant  $K_D = [Fab][Ag]/[Fab-Ag]$ , where  $[Fab]$ ,  $[Ag]$ , and  $[Fab-Ag]$  are the concentrations of the antibody Fab, antigen, and the complex, respectively (Azimzadeh and Van Regenmortel, 1990). In bivalent molecules interacting with binding partners that are tethered to a surface, the apparent affinity, or avidity, can be enhanced by multivalent binding. Such multivalent interactions are seen in many biological contexts including cell-cell communication, virus-host cell interactions, antibody-antigen interactions, and Fc receptor interactions with antigen-antibody complexes (Kiesling and Lamanna, 2003). Avidity effects benefit these interactions from both kinetic and thermodynamic standpoints. Binding bivalently to tethered binding partners is advantageous kinetically because if one arm dissociates, the likelihood of it finding its binding partner is greater due to the constraint of being tethered (Krishnamurthy et al., 2006). Avidity effects are also advantageous thermodynamically; whereas binding the first arm results in losses of translational and rotational degrees of freedom, the subsequent binding of the second arm incurs a

smaller entropic cost, thereby increasing the likelihood of the bivalent state (Krishnamurthy et al., 2006).

In the context of an IgG with two antigen-binding Fabs, the ability to bind bivalently to a virus is dependent on geometric factors such as the separation distances and orientations of tethered epitopes either on adjacent spikes during inter-spike crosslinking (Figure 1A, red box) or on individual spikes if intra-spike crosslinking can occur (Figure 1A, blue box) (Karush, 1976). Because the large distances between HIV-1 spikes makes inter-spike crosslinking unlikely, in this work, we focus exclusively on the latter mechanism of achieving bivalent binding. Although IgGs are too small to intra-spike crosslink (Figure 1A, gold box) (Klein, 2009; Wang et al., 2017), we previously engineered larger reagents (homo- and hetero-diFabs) that were designed to bind to a single Env, resulting in mean neutralization potency increases over a panel of HIV-1 strains (Galimidi et al., 2015). These diFab constructs were composed of two IgG Fabs joined by different lengths of double-stranded DNA (dsDNA), which served as both a rigid linker and a molecular ruler to probe the conformations of HIV-1 Env on virions (Galimidi et al., 2015) (Figure 1B). The dsDNA was flanked by two short single-stranded DNA (ssDNA) segments, where the primary differences between the two types of DNA is that dsDNA is more rigid and shorter (0.34 nm/bp) than the more flexible and longer ssDNA (0.64 nm/base).

In this work, we develop a statistical mechanical model that quantifies the potency of a diFab using its free energy of binding, which consists of two components. The first is the energy gained when each Fab attaches to Env, a quantity that only depends upon the identity of the Fab (the  $V_H$ ,  $V_L$ ,  $C_H1$ , and  $C_L$  segments in Figure 1B) and HIV-1 strain considered. The second component of the free energy is the entropic cost of binding either one or both Fabs to an Env, since a bound state restricts the number of possible conformations of a diFab, decreasing its likelihood. In our model, we enumerate the possible conformations using a random walk model for the diFab linker (the  $d$  bp dsDNA flanked by  $s$  bases ssDNA in Figure 1B).

Using our model, we can expand upon the earlier results of these synthetic diFab constructs and theoretically analyze whether changing the flexibility of the linker joining the two Fabs could also enhance neutralization potency. This enables us to compare a spectrum of possibilities from a rigid linker solely comprising dsDNA to a fully flexible linker composed of only ssDNA. We then generalize our model to a triFab design and demonstrate that simultaneously binding to three Env epitopes can greatly boost avidity. Insights from our synthetic constructs can be adapted to antibody design in other systems, in which length and rigidity of linkers in multivalent reagents must be balanced to elicit the most effective response.

## RESULTS

### Estimating the Parameters of diFab Binding from Crystal Structures

While HIV-1 Env fluctuates between multiple conformations, we assume that a diFab neutralizes the virus by binding to one specific state of Env at which the distance between the C-termini of the two Fabs (where the DNA is joined) is defined to be  $l_{\text{linker}}$  (Figure 2A). For example, the predicted distance based on

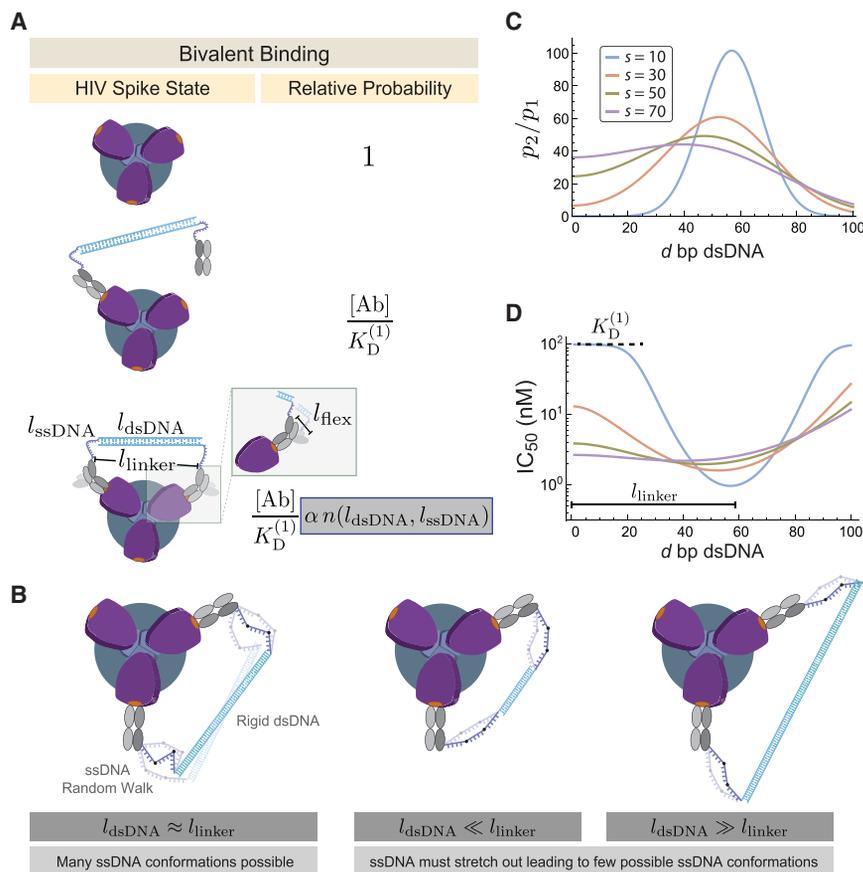
modeling adjacent 3BNC60 Fabs on a low-resolution open structure of an HIV-1 trimer (Merk and Subramaniam, 2013) was  $\approx 20$  nm. More recently, we used the 3BNC117 Fab-gp120 portion of a cryo-EM structure to measure the distance between adjacent Fab  $C_H1$  C-termini in the closed conformation of Env and then modeled a 3BNC60-gp120 protomer into three recent cryo-EM structures of Env trimers in different conformations: an open Env bound by the b12 bNAb, in which the coreceptor binding sites on the V3 loops are not exposed (Ozorowski et al., 2017); an open CD4-bound Env structure with exposed V3 loops (Ozorowski et al., 2017); and a partially open CD4-bound Env, in which the gp120 subunits adopted positions mid-way between closed and fully open (Wang et al., 2018). From these structures, we measured distances of 15.8, 20.3, 20.4, and 20.1 nm between C-termini of Fab  $C_H1$  domains modeled onto the closed, b12-bound open, CD4-bound open, and CD4-bound partially open Env conformations, respectively. Based on these values, we set  $l_{\text{linker}} = 20$  nm.

A further model parameter,  $l_{\text{flex}} = 1$  nm, shown in Figure 2A was included to account for the flexibility of the Fab. More precisely, this parameter accounts for variations in the distance between the C-termini of the two Fab  $C_H1$  domains to which the DNA was attached due to the following factors: (1) the Fab  $C_H1$ - $C_L$  domains can adopt different conformations with respect to  $V_H$ - $V_L$  (Stanfield et al., 2006) such that the locations of the  $C_H1$  C-terminus could shift by up to  $\approx 1$  nm; (2) residues C-terminal to  $C_H1$  residue 217 were found to be disordered in the 3BNC60 Fab structure (Scheid et al., 2011), thus the position of the  $C_H1$  residue to which the DNA was attached (Cys233) is uncertain within  $\approx 1$  nm; (3) the ssDNA is covalently linked to the  $C_H1$  residue Cys233 using an amine-to-sulfhydryl crosslinker, which exerts unknown effects on the length and the degree of flexibility between the ssDNA and Fab.

### Relating Neutralization to the Probability that an HIV-1 Spike Is Bound

To model diFab efficacy in terms of the properties of the linker, we first related the ability of a diFab to neutralize an HIV-1 virion to the probability that an Env spike on the surface of HIV-1 will be bound by an antibody. We assume that the spikes are sufficiently far apart to preclude inter-spike crosslinking (Figure 1A, red boxes) and focus exclusively on intra-spike crosslinking between the three identical sites on the Env homotrimer (Figure 1A, blue boxes). We further assume that viral infectivity varies linearly with the number of unbound Env, rising from zero (when all spikes are bound by diFabs) to maximum infectivity (when all spikes are unbound) as discussed in the STAR Methods section “% Neutralization and HIV-1 Env Binding Probability” (Klasse, 2007; Brandenburg et al., 2015).

Given these assumptions, the ability of a diFab to neutralize HIV-1 is proportional to the probability that at least one of the binding sites on an Env spike will be bound (see STAR Methods section titled “% Neutralization and HIV-1 Env Binding Probability”). For example, if each Env protein has a 75% chance to be bound by a diFab, an average of 75% of the spikes on each virion will be bound, and by the linearity assumption, the HIV-1 virions will be 75% neutralized. This enables us to relate the experimentally determinable % neutralization for diFabs to the theoretically tractable probability that a single Env spike will



**Figure 2. Modeling the Optimal Linker Length for a diFab**

(A) The energetic and entropic costs of monovalent binding depend upon the concentration  $[Ab]$  and dissociation constant  $K_D^{(1)}$  of a single Fab binding. The boost in bivalent binding is dictated by the geometric or avidity factors  $\alpha$  (which is the same for all diFabs) and  $n$  (which depends upon the length of the dsDNA and ssDNA, the optimal length  $l_{linker}$  of the linker between two bound Fabs, and the flexibility  $l_{flex}$  between the C<sub>H</sub>1-C<sub>L</sub> and V<sub>H</sub>-V<sub>L</sub> domains of a bound Fab). The ability to neutralize is given by the sum of relative probabilities for the monovalent and bivalent states divided by the sum over all states.

(B) The boost  $\alpha n$  of bivalently binding is computed by treating the ssDNA as a random walk and the dsDNA as a rigid rod (see STAR Methods section titled “Number of Microstates in the Model Including ssDNA”). An optimal linker matches the length of the dsDNA to the length between two bound Fabs ( $l_{dsDNA} \approx l_{linker}$ ), giving rise to many configurations for the bivalently bound state and increasing its likelihood.

(C and D) The relative probability that a diFab with  $d$  bp dsDNA and  $s$  bases ssDNA is bivalently ( $p_2$ ) versus monovalently ( $p_1$ ) bound (C) and the predicted  $IC_{50}$  values for these same constructs (D). Parameters used were  $l_{linker} = 20$  nm,  $l_{flex} = 1$  nm,  $\alpha = 5 \times 10^6$ , and  $K_D^{(1)} = 100$  nM.

be bound either monovalently or bivalently by a diFab. Avidity effects will allow an optimal diFab to bind more tightly to a spike, increasing the binding probability and the neutralization potency.

### The Avidity of a diFab Is Dictated by Its Linker Composition

To calculate the probability that any of the Fab binding sites on an HIV-1 spike are occupied, we compute the free energy (the energetic boost of the Fab-virus interaction and the entropic penalty of confining the number of linker states) for the three potential states of the spike, which represent a single diFab bound to zero, one, or two binding sites (Figure 2A). The entropy of the linker was characterized by treating the dsDNA as a 1D rigid rod and the ssDNA as a random walk. The former assumption is valid provided the dsDNA in each linker is less than the 150 bp persistence length of dsDNA (Brunet et al., 2015), a reasonable restriction given that only 60 bp dsDNA are required to span  $l_{linker} = 20$  nm. Free ssDNA is flexible with a persistence length  $\xi_{ssDNA} = 1.5$  nm ( $\approx 2.3$  bases) (Tinland et al., 1997; Ambia-Garrido et al., 2010) that we analyze using the ideal chain polymer physics model (Doi and Edwards, 1986).

When a diFab transitions from the unbound state (with probability  $p_0$ ) to a monovalently bound state (with probability  $p_1$ ), it loses translational and rotational entropy but gains favorable binding energy (Phillips et al., 2013) leading to the relative probability

$$\frac{p_1}{p_0} = \frac{[Ab]}{K_D^{(1)}}, \quad (\text{Equation 1})$$

where  $[Ab]$  is the concentration of the diFab and  $K_D^{(1)}$  is the equilibrium dissociation constant of the first diFab arm binding to Env.  $K_D^{(1)}$  can be experimentally determined as the  $IC_{50}$  (concentration of antibody capable of neutralizing 50% of the virus) of a Fab neutralization profile, with a smaller  $IC_{50}$  signifying a more potent antibody. We use the typical value  $K_D^{(1)} = 100$  nM reported for a CD4-binding site bNAb dissociating from a soluble, trimeric HIV-1 Env (Lyumkis et al., 2013). Importantly, the transition from an unbound to a monovalently bound diFab is independent of the amount of dsDNA and ssDNA in the diFab’s linker.

We now turn to the transition from a monovalently bound diFab to a bivalently bound diFab. The ability to simultaneously bind two epitopes depends on the linker composition (the quantity of dsDNA and ssDNA), since the distance  $l_{linker}$  between the C-termini of bivalently bound Fabs must be spanned by the rigid dsDNA segments as well as the two flanking ssDNA strands (Figure 2A). More precisely, we enumerate the configurations of the ssDNA random walk and the dsDNA in the bivalently bound state (Figure 2B), permitting us to compute the entropic cost of bivalent binding (see STAR Methods section titled “Number of Microstates in the Model Including ssDNA”). Within this framework, the probability of the bivalently bound state ( $p_2$ ) compared

to the monovalently bound state is given by the product of a constant  $\alpha$  that is independent of the linker composition and a term  $n(l_{\text{dsDNA}}, l_{\text{ssDNA}})$  that depends on the quantity of dsDNA and ssDNA, namely,

$$\frac{p_2}{p_1} = \alpha n$$

$$= \frac{\alpha l_{\text{flex}}^2}{l_{\text{linker}} l_{\text{dsDNA}}} \sqrt{\frac{3\xi_{\text{ssDNA}}}{8\pi^3 l_{\text{ssDNA}}}} e^{-\frac{3(l_{\text{linker}}^2 + l_{\text{dsDNA}}^2)}{8(l_{\text{ssDNA}} + l_{\text{ssDNA}})}} \sinh\left(\frac{3l_{\text{linker}} l_{\text{dsDNA}}}{4l_{\text{ssDNA}} \xi_{\text{ssDNA}}}\right)$$

(Equation 2)

where  $l_{\text{dsDNA}} = d(0.34 \text{ nm/bp})$  and  $l_{\text{ssDNA}} = s(0.64 \text{ nm/base})$  represent the length of  $d$  dsDNA base pairs and  $s$  ssDNA bases in the linker, respectively.

Equations 1 and 2, together with the normalization condition  $p_0 + p_1 + p_2 = 1$ , enable us to compute the probability that a diFab will neutralize a virion,

$$p_1 + p_2 = \frac{\frac{[\text{Ab}]}{K_D^{(1)}} + \frac{[\text{Ab}]}{K_D^{(1)}} \alpha n}{1 + \frac{[\text{Ab}]}{K_D^{(1)}} + \frac{[\text{Ab}]}{K_D^{(1)}} \alpha n},$$

(Equation 3)

from which we can write the concentration of 50% inhibition,

$$IC_{50}^{\text{diFab}} = \frac{K_D^{(1)}}{1 + \alpha n}.$$

(Equation 4)

The value of  $\alpha = 5 \times 10^6$  was calibrated from previous measurements where a diFab with  $d=62$  bp and  $s = 12$  bases neutralized HIV-1 approximately 100-fold better than the Fab alone (Galimidi et al., 2015). Figure 2C compares the probability that a diFab will be bivalently bound ( $p_2$ ) rather than monovalently bound ( $p_1$ ) for different linkers. The model shows that bivalent binding is most likely when  $l_{\text{linker}} \approx l_{\text{dsDNA}}$ , when the rigid dsDNA approximately spans the length between the two bound Fabs. This peak shifts leftwards with the root-mean-squared length of the flexible ssDNA, demonstrating that ssDNA can make up for dsDNA that is slightly too short or too long, provided the flexibility of the ssDNA is taken into account. Figure 2D shows the corresponding  $IC_{50}$ s for these constructs. Although adding more flexible ssDNA leads to a broader segment of dsDNA lengths capable of enhanced neutralization, the optimal diFab potency is achieved by including less ssDNA and maximizing the rigidity of the linker.

### The diFab Model Allows Bivalent Binding Only When the dsDNA Length Is Approximately Equal to the Length of the Linker It Spans

To gain a qualitative understanding of our results, we examined Equation 4 in two limits: near the optimal geometry  $l_{\text{dsDNA}} \approx l_{\text{linker}}$  where the ability to bind bivalently is maximum and far from the optimal limit when the diFab is too short or too long to permit bivalent binding through intra-spike crosslinking.

Near the optimal geometry, HIV-1 neutralization occurs predominantly from the bivalently bound configuration rather than the monovalent state,  $\frac{p_2}{p_1} = \alpha n \gg 1$ . Hence, the system is well approximated with each spike either being unbound or bivalently

bound, with the dissociation constant  $K_D^{(1)}$  for a single Fab boosted by the avidity factor  $\alpha n$ , namely,

$$IC_{50}^{\text{diFab}} \approx \frac{K_D^{(1)}}{\alpha n} \quad (\text{near optimal geometry}) \quad (\text{Equation 5})$$

If the potency of diFab 1 is  $IC_{50}^{(1)}$  and the potency of diFab 2 with a different linker is  $IC_{50}^{(2)}$ , the latter diFab's potency will be shifted relative to the former by the ratio of  $n$  factors, namely,

$$\frac{IC_{50}^{(2)}}{IC_{50}^{(1)}} = \frac{n^{(1)}}{n^{(2)}}.$$

(Equation 6)

In other words, the relative potency of both diFabs is determined solely by the entropy, rather than the energy, of the linker when bivalently bound.

In the opposite regime where the diFab linker is too small ( $l_{\text{dsDNA}} + l_{\text{ssDNA}} < l_{\text{linker}}$ ) or too large ( $l_{\text{dsDNA}} - l_{\text{ssDNA}} > l_{\text{linker}}$ ), the diFab loses the ability to bind bivalently and the  $IC_{50}$  attains a constant value

$$IC_{50} \approx K_D^{(1)} \quad (\text{far from optimal geometry}) \quad (\text{Equation 7})$$

shown as a black dashed line in Figure 2D.

### Avidity Is Further Enhanced Using a triFab that Binds Three Env Sites

While the diFab constructs were inspired by two-armed IgG antibodies, a triFab construct that could simultaneously bind to three Env epitopes (Figure 1B) should exhibit even greater avidity and hence neutralize HIV-1 more potently. For simplicity, we assume that both dsDNA segments have the same length, as do all ssDNA segments. As derived in the section titled "Generalizing the Model to GappedFabs and TriFabs" of the STAR Methods, the  $IC_{50}$  of such a construct is given by

$$IC_{50}^{\text{triFab}} = \frac{K_D^{(1)}}{1 + \alpha n + \alpha^2 n^2}.$$

(Equation 8)

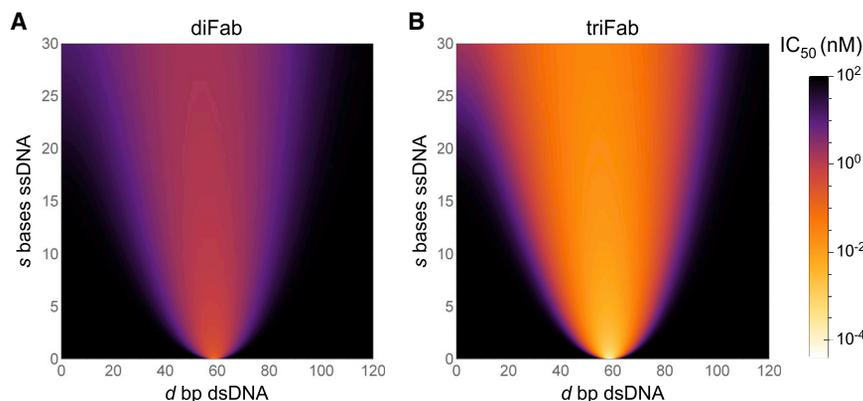
Figures 3A and 3B compare the neutralization potencies across the design space of diFabs and triFabs, demonstrating that joining three Fabs can result in  $IC_{50}$ s far smaller than what is possible for even the theoretically optimal diFab design.

The neutralization potency of a triFab near its optimal geometry ( $\alpha n \gg 1$ ) is dictated purely by the trivalently bound state,

$$IC_{50}^{\text{triFab}} \approx \frac{K_D^{(1)}}{\alpha^2 n^2} \quad (\text{near optimal geometry}) \quad (\text{Equation 9})$$

Note that the boost in avidity in going from a Fab to a diFab is equivalent to the boost between a diFab and triFab. Since diFabs have been shown to achieve a 100–1,000 fold decrease in  $IC_{50}$  over a Fab, a triFab should be able to achieve a  $10^4$ – $10^6$ -fold decrease in  $IC_{50}$  relative to the Fab, providing a powerful framework with which to achieve very high neutralization.

To put these results into perspective, potent IgGs commonly have  $IC_{50}$  values on the order of 1 nM across a panel of HIV-1 viruses (Galimidi et al., 2015). While the effects of an escape mutation can be highly heterogeneous (some shifting  $IC_{50}$  by



**Figure 3. Neutralization Potency of Different diFab and triFab Designs**

(A and B) Predicted  $IC_{50}$ s for (A) diFab or (B) triFab constructs with linkers composed of  $d$  bp dsDNA and  $s$  bases ssDNA segments. In both cases, optimal neutralization occurs when the linker matches the length between Fab binding sites ( $d = (3 \text{ bp/nm})l_{\text{linker}} = 60 \text{ bp}$ ) and has as little ssDNA as possible. The triFab can achieve significantly enhanced potency over the diFab because of its additional avidity.

as little as 2-fold while others can increase  $IC_{50}$  by more than 100-fold [Falkowska et al., 2014; Dingens et al., 2017; Lee et al., 2017]), escape mutations commonly increase  $IC_{50}$  by roughly 10x. Assuming this rule of thumb, an optimal diFab can withstand 2 escape mutations and remain as efficacious as a potent IgG while an optimal triFab can tolerate 4 such mutations.

To further enhance potency, more than three Fabs could be joined together into a multiFab. By joining multiple types of Fabs that bind to different epitopes (e.g., the broadly neutralizing CD4-binding 3BNC117 with a V1V2-binding PG16), multiFabs may enhance the potency of neutralization as well as help combat HIV-1 heterogeneity as has been seen in combination influenza antibodies [Laursen et al., 2018] and HIV-1 mosaic vaccines [Bricault et al., 2019].

### An Optimal Linker Is Maximally Rigid and Perfectly Spans the Distance between Env Epitopes

Counter to what might be intuitively expected, adding flexibility to a linker by increasing the number of ssDNA bases need not improve the diFab's neutralization potential. This effect arises because flexible ssDNA has a large number of degrees of freedom that are constrained when the diFab bivalently binds, leading to a larger entropic penalty (or smaller boost in avidity). Indeed, the full diFab and triFab design space shown in Figure 3 demonstrates that an optimal construct is a perfectly rigid linker composed of only dsDNA whose length matches the distance between HIV-1 epitopes ( $l_{\text{dsDNA}} \approx l_{\text{linker}}$ ). However, we caution that such diFabs may not operate optimally for both experimental and theoretical reasons including: (1) the range of dsDNA lengths that permit bivalent binding narrows as the amount of ssDNA decrease, and in the extreme limit of a linker with no ssDNA, being slightly too short or too long by as little as a few base pairs may preclude bivalent binding; (2) charge interactions between the dsDNA and either the Fab or Env may disrupt diFab functionality; (3) a lack of sufficient flexibility at the Fab-dsDNA junction may preclude bivalent binding; and (4) in the limit of a rigid linker (ssDNA  $\leq 10$  bases) the ideal chain model breaks down and the bending of ssDNA must be accounted for (e.g., via the worm-like chain model), which may result in higher  $IC_{50}$ s than predicted by our model.

Because of these considerations, it is worthwhile to use a diFab whose dsDNA length  $l_{\text{dsDNA}}$  matches the distance between Fab epitopes ( $l_{\text{linker}}$ ) found in crystal structures, and to flank

this dsDNA with short segments of ssDNA ( $\approx 10$  bases) to buffer against uncertainties in the measurements. As seen

in Figure 3, such constructs lie within a wide basin with strong neutralization potential. Of note, this strategy far surpasses a purely flexible linker ( $l_{\text{ssDNA}} \approx l_{\text{linker}}$ ,  $l_{\text{dsDNA}} = 0$ ), which cannot bivalently bind because of the large entropic cost. Other methods of introducing flexibility, such as introducing ssDNA breaks in the dsDNA linker, are also predicted to increase the entropic cost of bivalent binding and hence increase the  $IC_{50}$  (see STAR Methods section titled "Generalizing the Model to GappedFabs and TriFabs").

## DISCUSSION

The low density of Env spikes on HIV-1 potentially enables the virus to mitigate the host antibody response by hindering IgGs from using both antigen-binding Fabs to bind bivalently, thereby expanding the range of HIV-1 mutations permitting antibody evasion [Klein and Bjorkman, 2010; Galimidi et al., 2015]. Indeed, a mutant simian immunodeficiency virus (SIV) with a higher number of Env spikes reverted to its normal spike count of  $\approx 14$  when propagated in non-human primate hosts [Zingler and Littman, 1993]. This suggests that while HIV-1 may be more infectious with more Env trimers [Layne et al., 1990], the immune system applies selective pressure that keeps the Env spike count per virion low, presumably to prevent anti-HIV-1 IgGs from utilizing avidity effects to counter the lower intrinsic Fab-Env affinities that result from rapid mutation of Env. Antibody isotypes, such as dimeric IgA or pentameric IgM, have increased valencies (four and ten Fabs, respectively) compared to the two Fabs of an IgG, thus allowing for increased avidity effects during antibody binding to a pathogen. However, most of the neutralizing activity in the sera of HIV-1-positive individuals is attributed to IgGs [Tomaras et al., 2008; Scheid et al., 2009], and converting an anti-HIV-1 IgG antibody to an IgA or IgM has minimal effects on potency in standard neutralization assays [Wolbank et al., 2003; Kunert et al., 2004], possibly because there are so few Env spaced sufficiently far apart that these other antibody classes are also forced to bind monovalently.

Bivalent binding to single Env trimer (intra-spike crosslinking) is another way to utilize avidity effects to counteract the low spike density of HIV-1. Although the architecture of Env trimers prohibits this mode of binding for conventional, host-derived IgGs [Klein, 2009; Wang et al., 2017], we analyzed how synthetic diFabs (Fabs from a neutralizing anti-HIV-1 IgG joined by a linker

containing rigid dsDNA flanked by flexible ssDNA shown in [Figure 1B](#)) could be designed to achieve optimal intra-spike crosslinking.

HIV-1 Env trimers adopt multiple conformations on virions ([Liu et al., 2008](#); [Munro et al., 2014](#)) and in the soluble native-like forms used for structural studies ([Ward and Wilson, 2017](#)). For example, binding of the host CD4 receptor induces outward displacements of the three Env gp120 subunits, resulting in an open conformation in which the coreceptor binding sites on the trimer apex V3 loops are exposed ([Harris et al., 2011](#); [Wang et al., 2016, 2018](#); [Ozorowski et al., 2017](#)) and that rearranges further upon coreceptor binding and subsequent membrane fusion. We measured the distances between adjacent 3BNC117 epitopes in a new cryo-EM structure ([Lee et al., 2017](#)) and estimated the average position of the C-terminal C<sub>H</sub>1 domain residue to which the DNA of a 3BNC117 diFab would be covalently attached ([Stanfield et al., 2006](#)). Based on these measurements, we assumed that a diFab neutralized HIV-1 when the linker spans a length  $l_{\text{linker}} = 20$  nm. We further assumed that the Fab C<sub>H</sub>1-C<sub>L</sub> domains can stretch relative to the V<sub>H</sub>-V<sub>L</sub> domains by  $l_{\text{flex}} = 1$  nm (see [Figure 2A](#)). Lastly, we considered a Fab whose ability to dissociate (as given by the midpoint of a neutralization assay) equals to  $K_D^{(1)} = 100$  nM. Each of these values can be readily adapted to other HIV-1 strains and Fabs.

With these parameters in hand, we developed a statistical mechanics-based model to predict the neutralization of a diFab whose linker is composed of  $d$  base pairs dsDNA and  $s$  bases ssDNA ([Figure 1B](#)), enabling us to tune both the length and rigidity of the linker. By assuming that (i) each homotrimeric spike is unable to help infect a host cell when any one of its three epitopes are bound by Fab and (ii) that the infectivity of a virion varies linearly with the number of unbound Env, we showed that the neutralization of a virion is proportional to the probability that any single Env protein is bound by a Fab (see [STAR Methods](#) section titled “% Neutralization and HIV-1 Env Binding Probability”). This framework enabled us to translate the linker-dependent entropy and energy of binding to an HIV-1 Env trimer into the predicted neutralization potency for each diFab.

It is worthwhile to point out several factors that the model neglects. First, the model does not consider potential, but presumably rare, diFab binding between adjacent Env trimers. Second, our model assumed that the % neutralization of HIV-1 decreases linearly with the number of unbound Env trimers. While such a linear relationship was observed when less than half of HIV-1 spikes were bound ([Layne et al., 1990](#)), it may break down if, for example, at least 2–3 unbound Env trimers are needed to infect a cell. We also assumed that each virion had exactly 14 spikes, neglecting the relatively small observed spike number variations ([Chertova et al., 2002](#); [Zhu et al., 2003, 2006](#); [Liu et al., 2008](#)). However, relaxing these assumptions minimally changes the neutralization curves (which may become slightly sharper, as shown in the [STAR Methods](#) section titled “% Neutralization and HIV-1 Env Binding Probability”), suggesting that our results are robust and should apply to more general descriptions of HIV-1 neutralization.

We determined that an optimal linker will maximize its rigidity, trading flexible ssDNA for rigid dsDNA. The larger flexibility of ssDNA implies that there will be a higher entropic penalty for

bivalently binding, thereby resulting in a worse (larger) IC<sub>50</sub>. This general statement applies to all forms of increased flexibility including: (1) trading dsDNA for an equivalent length of ssDNA, (2) adding ssDNA without decreasing the length of dsDNA, or (3) introducing ssDNA gaps in the dsDNA (see [STAR Methods](#) section titled “Generalizing the Model to GappedFabs and TriFabs”).

In addition to tuning the length of dsDNA and ssDNA in the linker, another biologically inspired approach to further increase avidity is to construct multiFabs that target more than two epitopes. As has been seen in other biological systems ([Kitov and Bundle, 2003](#); [Liese and Netz, 2015](#); [Yan et al., 2018](#)), higher valencies can elicit tighter binding. Hence, a triFab that allows three Fabs to simultaneously bind to an HIV-1 Env trimer is predicted to have a lower IC<sub>50</sub> than an optimal diFab; indeed, our model predicts that the boost from avidity of an optimal diFab over a Fab is equivalent to the boost of an optimal triFab over an optimal diFab, and hence we expect that triFabs should be able to achieve IC<sub>50</sub>s 10<sup>4</sup>–10<sup>6</sup> fold smaller than Fabs. MultiFabs that link together Fabs targeting different epitopes (e.g., the broadly neutralizing CD4-binding site antibody 3BNC117 and the V2-binding PG16 antibody) ([Bricault et al., 2019](#)) could better combat the heterogeneity of HIV-1 strains, providing guidance for constructing optimal anti-HIV-1 therapeutics that remain potent against HIV-1 in the face of the Env mutations arising during HIV-1 replication.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **LEAD CONTACT AND MATERIALS AVAILABILITY**
  - Materials Availability Statement
- **METHOD DETAILS**
  - % Neutralization and HIV-1 Env Binding Probability
  - Number of Microstates in the Model including ssDNA
  - Generalizing the Model to GappedFabs and TriFabs
- **DATA AND CODE AVAILABILITY**

## SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.cels.2019.09.007>.

## ACKNOWLEDGMENTS

We thank Anthony Bartolotta, Justin Bois, Jim Eisenstein, Vahe Galstyan, Peng He, Willem Kegel, David Hsieh, Giacomo Koszegi, Pankaj Mehta, Jiseon Min, Olexei Motrunich, Noah Olsman, Vahe Singh, and Richard Zhu for useful discussions; Christopher Barnes for measuring modeled 3BNC60-Env complexes; and Marta Murphy for help preparing figures. This research was supported by NIH NIAID grants 1R01AI129784 and H1VRAD P01 AI100148 (P.J.B.), the Bill and Melinda Gates Foundation Collaboration for AIDS Vaccine Discovery grant 1040753 (P.J.B.), La Fondation Pierre-Gilles de Gennes (R.P.), the Rosen Center at Caltech (R.P.), 1R35 GM118043-01 Maximizing Investigators’ Research Award (MIRA), United States, and R01 GM085286 (R.P.), and a Caltech-COH Biomedical Research Initiative (P.J.B.). We thank the Burroughs Wellcome Fund for their support through the Career Award at the Scientific Interface (S.Y.) as well as for the Physiology Course at the Marine Biological Laboratory where part of this work was done.

## AUTHOR CONTRIBUTIONS

T.E., A.P.W., R.P., and P.J.B. conceived the project. T.E., S.Y., and R.P. developed the model and performed analyses. T.E., R.P., and P.J.B. wrote the paper with input from other authors.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 22, 2019

Revised: September 10, 2019

Accepted: September 17, 2019

Published: October 23, 2019

## REFERENCES

- Ambia-Garrido, J., Vainrub, A., and Pettitt, B.M. (2010). A model for Structure and Thermodynamics of ssDNA and dsDNA Near a Surface: a Coarse Grained Approach. *Comput. Phys. Commun.* *181*, 2001–2007.
- Azimzadeh, A., and Van Regenmortel, M.H. (1990). Antibody affinity measurements. *J. Mol. Recognit.* *3*, 108–116.
- Baba, T.W., Liska, V., Hofmann-Lehmann, R., Vlasak, J., Xu, W., Ayehunie, S., Cavacini, L.A., Posner, M.R., Katinger, H., Stiegler, G., et al. (2000). Human neutralizing monoclonal antibodies of the IgG1 subtype protect against mucosal simian-human immunodeficiency virus infection. *Nat. Med.* *6*, 200–206.
- Berg, H.C. (1993). *Random walks in biology* (Princeton University Press).
- Brandenberg, O.F., Magnus, C., Rusert, P., Günthard, H.F., Regoes, R.R., and Trkola, A. (2017). Predicting HIV-1 transmission and antibody neutralization efficacy in vivo from stoichiometric parameters. *PLoS Pathog.* *13*, e1006313.
- Brandenberg, O.F., Magnus, C., Rusert, P., Regoes, R.R., and Trkola, A. (2015). Different infectivity of HIV-1 strains is linked to number of envelope trimers required for entry. *PLoS Pathog.* *11*, e1004595.
- Bricault, C.A., Yusim, K., Seaman, M.S., Yoon, H., Theiler, J., Giorgi, E.E., Wagh, K., Theiler, M., Hraber, P., Macke, J.P., et al. (2019). HIV-1 neutralizing antibody signatures and application to epitope-targeted vaccine design. *Cell Host Microbe* *25*, 59–72.
- Brunet, A., Tardin, C., Salomé, L., Rousseau, P., Destainville, N., and Manghi, M. (2015). Dependence of DNA persistence length on ionic strength of solutions with monovalent and divalent salts: A joint theory–experiment study. *Macromolecules* *48*, 3641–3652.
- Caskey, M., Klein, F., Lorenzi, J.C., Seaman, M.S., West, A.P., Jr., Buckley, N., Kremer, G., Nogueira, L., Braunschweig, M., Scheid, J.F., et al. (2015). Viraemia suppressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117. *Nature* *522*, 487–491.
- Caskey, M., Schoofs, T., Gruell, H., Settler, A., Karagounis, T., Kreider, E.F., Murrell, B., Pfeifer, N., Nogueira, L., Oliveira, T.Y., et al. (2017). Antibody 10–1074 suppresses viremia in HIV-1-infected individuals. *Nat. Med.* *23*, 185–191.
- Chertova, E., Bess, J.W., Jr., Crise, B.J., Sowder II, R.C., Schaden, T.M., Hilburn, J.M., Hoxie, J.A., Benveniste, R.E., Lifson, J.D., Henderson, L.E., et al. (2002). Envelope glycoprotein incorporation, not shedding of surface envelope glycoprotein (gp120/SU), is the primary determinant of SU content of purified human immunodeficiency virus type 1 and simian immunodeficiency virus. *J. Virol.* *76*, 5315–5325.
- Dingens, A.S., Haddox, H.K., Overbaugh, J., and Bloom, J.D. (2017). Comprehensive mapping of HIV-1 escape from a broadly neutralizing antibody. *Cell Host Microbe* *21*, 777–787.e4.
- Doi, M., and Edwards, S.F. (1986). *Static properties of polymers* (Oxford University Press).
- Doria-Rose, N.A., Schramm, C.A., Gorman, J., Moore, P.L., Bhiman, J.N., DeKosky, B.J., Ernandes, M.J., Georgiev, I.S., Kim, H.J., Pancera, M., et al. (2014). Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. *Nature* *509*, 55–62.
- Eisen, H.N., and Siskind, G.W. (1964). Variations in affinities of antibodies during the immune response. *Biochemistry* *3*, 996–1008.
- Escolano, A., Dosenovic, P., and Nussenzweig, M.C. (2017). Progress toward active or passive HIV-1 vaccination. *J. Exp. Med.* *214*, 3–16.
- Falkowska, E., Le, K.M., Ramos, A., Doores, K.J., Lee, J.H., Blattner, C., Ramirez, A., Derking, R., van Gils, M.J., Liang, C.H., et al. (2014). Broadly neutralizing HIV antibodies define a glycan-dependent epitope on the prefusion conformation of gp41 on cleaved envelope trimers. *Immunity* *40*, 657–668.
- Fauci, A.S. (2008). 25 years of HIV. *Nature* *453*, 289–290.
- Galimidi, R.P., Klein, J.S., Politzer, M.S., Bai, S., Seaman, M.S., Nussenzweig, M.C., West, A.P., Jr., and Bjorkman, P.J. (2015). Intra-spike crosslinking overcomes antibody evasion by HIV-1. *Cell* *160*, 433–446.
- Harris, A., Borgnia, M.J., Shi, D., Bartesaghi, A., He, H., Pejchal, R., Kang, Y.K., Depetris, R., Marozsan, A.J., Sanders, R.W., et al. (2011). Trimeric HIV-1 glycoprotein gp140 immunogens and native HIV-1 envelope glycoproteins display the same closed and open quaternary molecular architectures. *Proc. Natl. Acad. Sci. USA* *108*, 11440–11445.
- Hessell, A.J., Hangartner, L., Hunter, M., Havenith, C.E., Beurskens, F.J., Bakker, J.M., Lanigan, C.M., Landucci, G., Forthal, D.N., Parren, P.W., et al. (2007). Fc receptor but not complement binding is important in antibody protection against HIV. *Nature* *449*, 101–104.
- Hessell, A.J., Rakasz, E.G., Poignard, P., Hangartner, L., Landucci, G., Forthal, D.N., Koff, W.C., Watkins, D.I., and Burton, D.R. (2009). Broadly neutralizing human anti-HIV antibody 2G12 is effective in protection against mucosal SHIV challenge even at low serum neutralizing titers. *PLoS Pathog.* *5*, e1000433.
- Karush, F. (1976). Multivalent binding and functional affinity. *Contemp. Top. Mol. Immunol.* *5*, 217–228.
- Kiessling, L.L., and Lamanna, A.C. (2003). Multivalency in biological systems (Springer).
- Kitov, P.I., and Bundle, D.R. (2003). On the nature of the multivalency effect: a thermodynamic model. *J. Am. Chem. Soc.* *125*, 16271–16284.
- Klasse, P.J. (2007). Modeling how many envelope glycoprotein trimers per virion participate in human immunodeficiency virus infectivity and its neutralization by antibody. *Virology* *369*, 245–262.
- Klasse, P.J. (2012). The molecular basis of HIV entry. *Cell. Microbiol.* *14*, 1183–1192.
- Klein, F., Halper-Stromberg, A., Horwitz, J.A., Gruell, H., Scheid, J.F., Bournazos, S., Mouquet, H., Spatz, L.A., Diskin, R., Abadir, A., et al. (2012). HIV therapy by a combination of broadly neutralizing antibodies in humanized mice. *Nature* *492*, 118–122.
- Klein, F., Mouquet, H., Dosenovic, P., Scheid, J.F., Scharf, L., and Nussenzweig, M.C. (2013). Antibodies in HIV-1 vaccine development and therapy. *Science* *341*, 1199–1204.
- Klein, F., Nogueira, L., Nishimura, Y., Phad, G., West, A.P., Jr., Halper-Stromberg, A., Horwitz, J.A., Gazumyan, A., Liu, C., Eisenreich, T.R., et al. (2014). Enhanced HIV-1 immunotherapy by commonly arising antibodies that target virus escape variants. *J. Exp. Med.* *211*, 2361–2372.
- Klein, J.S. (2009). Investigations in the design and characterization of HIV-1 neutralizing molecules, *Physiol.D* (California Institute of Technology).
- Klein, J.S., and Bjorkman, P.J. (2010). Few and far between: how HIV may be evading antibody avidity. *PLoS Pathog.* *6*, e1000908.
- Krishnamurthy, V.M., Estroff, L.A., and Whitesides, G.M. (2006). *Multivalency in ligand design* (Wiley-VCH Verlag GmbH & Co. KGaA).
- Kunert, R., Wolbank, S., Stiegler, G., Weik, R., and Katinger, H. (2004). Characterization of molecular features, antigen-binding, and in vitro properties of IgG and IgM variants of 4E10, an anti-HIV type 1 neutralizing monoclonal antibody. *AIDS Res. Hum. Retroviruses* *20*, 755–762.
- Laursen, N.S., Friesen, R.H.E., Zhu, X., Jongeneelen, M., Blokland, S., Vermond, J., van Eijgen, A., Tang, C., van Diepen, H., Obmolova, G., et al. (2018). Universal protection against influenza infection by a multidomain antibody to influenza hemagglutinin. *Science* *362*, 598–602.
- Layne, S.P., Merges, M.J., Dembo, M., Spouge, J.L., Conley, S.R., Moore, J.P., Raina, J.L., Renz, H., Gelderblom, H.R., and Nara, P.L. (1992). Factors

- underlying spontaneous inactivation and susceptibility to neutralization of human immunodeficiency virus. *Virology* 189, 695–714.
- Layne, S.P., Merges, M.J., Dembo, M., Spouge, J.L., and Nara, P.L. (1990). HIV requires multiple gp120 molecules for CD4-mediated infection. *Nature* 346, 277–279.
- 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). A Broadly Neutralizing antibody Targets the Dynamic HIV Envelope trimer Apex via a Long, Rigidified, and Anionic beta-Hairpin Structure. *Immunity* 46, 690–702.
- Liao, H.X., Lynch, R., Zhou, T., Gao, F., Alam, S.M., Boyd, S.D., Fire, A.Z., Roskin, K.M., Schramm, C.A., Zhang, Z., et al. (2013). Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature* 496, 469–476.
- Liese, S., and Netz, R.R. (2015). Influence of length and flexibility of spacers on the binding affinity of divalent ligands. *Beilstein J. Org. Chem.* 11, 804–816.
- Liu, J., Bartesaghi, A., Borgnia, M.J., Sapiro, G., and Subramaniam, S. (2008). Molecular architecture of native HIV-1 gp120 trimers. *Nature* 455, 109–113.
- Lynch, R.M., Boritz, E., Coates, E.E., DeZure, A., Madden, P., Costner, P., Enama, M.E., Plummer, S., Holman, L., Hendel, C.S., et al. (2015). Virologic effects of broadly neutralizing antibody VRC01 administration during chronic HIV-1 infection. *Sci. Transl. Med.* 7, 319ra206.
- Lyumkis, D., Julien, J.P., de Val, N., Cupo, A., Potter, C.S., Klasse, P.J., Burton, D.R., Sanders, R.W., Moore, J.P., Carragher, B., et al. (2013). Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. *Science* 342, 1484–1490.
- Mascola, J.R., Stiegler, G., VanCott, T.C., Katinger, H., Carpenter, C.B., Hanson, C.E., Beary, H., Hayes, D., Frankel, S.S., Bix, D.L., et al. (2000). Protection of macaques against vaginal transmission of a pathogenic HIV-1/SIV chimeric virus by passive infusion of neutralizing antibodies. *Nat. Med.* 6, 207–210.
- McCoy, L.E., and Burton, D.R. (2017). Identification and specificity of broadly neutralizing antibodies against HIV. *Immunol. Rev.* 275, 11–20.
- Merk, A., and Subramaniam, S. (2013). HIV-1 envelope glycoprotein structure. *Curr. Opin. Struct. Biol.* 23, 268–276.
- Montefiori, D.C. (2005). Evaluating neutralizing antibodies against HIV, SIV, and SHIV in luciferase reporter gene assays. *Curr. Protoc. Immunol.* 12, 11.
- Moore, P.L., Gray, E.S., Wibmer, C.K., Bhiman, J.N., Nonyane, M., Sheward, D.J., Hermanus, T., Bajimaya, S., Tumba, N.L., Abrahams, M.R., et al. (2012). Evolution of an HIV glycan-dependent broadly neutralizing antibody epitope through immune escape. *Nat. Med.* 18, 1688–1692.
- Munro, J.B., Gorman, J., Ma, X., Zhou, Z., Arthos, J., Burton, D.R., Koff, W.C., Courter, J.R., Smith, A.B., 3rd, Kwong, P.D., et al. (2014). Conformational dynamics of single HIV-1 envelope trimers on the surface of native virions. *Science* 346, 759–763.
- Ozorowski, G., Pallesen, J., de Val, N., Lyumkis, D., Cottrell, C.A., Torres, J.L., Copps, J., Stanfield, R.L., Cupo, A., Pugach, P., et al. (2017). Open and closed structures reveal allostery and pliability in the HIV-1 envelope spike. *Nature* 547, 360–363.
- Patel, P., Borkowf, C.B., Brooks, J.T., Lasry, A., Lansky, A., and Mermin, J. (2014). Estimating per-act HIV transmission risk: a systematic review. *AIDS* 28, 1509–1519.
- Phillips, R., Kondev, J., Theriot, J., and Garcia, H.G. (2013). *Physical Biology of the Cell* (Garland Science).
- Saphire, E.O., Parren, P.W., Pantophlet, R., Zwick, M.B., Morris, G.M., Rudd, P.M., Dwek, R.A., Stanfield, R.L., Burton, D.R., and Wilson, I.A. (2001). Crystal structure of a neutralizing human IgG against HIV-1: a template for vaccine design. *Science* 293, 1155–1159.
- Scheid, J.F., Mouquet, H., Feldhahn, N., Seaman, M.S., Velinzon, K., Pietzsch, J., Ott, R.G., Anthony, R.M., Zebroski, H., Hurley, A., et al. (2009). Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. *Nature* 458, 636–640.
- Scheid, J.F., Mouquet, H., Ueberheide, B., Diskin, R., Klein, F., Oliveira, T.Y., Pietzsch, J., Fenyo, D., Abadir, A., Velinzon, K., et al. (2011). Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science* 333, 1633–1637.
- Schoofs, T., Klein, F., Braunschweig, M., Kreider, E.F., Feldmann, A., Nogueira, L., Oliveira, T., Lorenzi, J.C., Parrish, E.H., Learn, G.H., et al. (2016). HIV-1 immunotherapy with monoclonal antibody 3BNC117 elicits. *Science* 352, 997–1001.
- Stanfield, R.L., Zemla, A., Wilson, I.A., and Rupp, B. (2006). Antibody elbow angles are influenced by their light chain class. *J. Mol. Biol.* 357, 1566–1574.
- Tinland, B., Pluen, A., Sturm, J., and Weill, G. (1997). Persistence length of single-stranded DNA. *Macromolecules* 30, 5763–5765.
- Tomaras, G.D., Yates, N.L., Liu, P., Qin, L., Fouda, G.G., Chavez, L.L., Decamp, A.C., Parks, R.J., Ashley, V.C., Lucas, J.T., et al. (2008). Initial B-cell responses to transmitted human immunodeficiency virus type 1: virion-binding immunoglobulin M (IgM) and IgG antibodies followed by plasma anti-gp41 antibodies with ineffective control of initial viremia. *J. Virol.* 82, 12449–12463.
- (2019). UNAIDS Fact Sheet.
- Wang, H., Barnes, C.O., Yang, Z., Nussenzweig, M.C., and Bjorkman, P.J. (2018). Partially open HIV-1 envelope structures exhibit conformational changes relevant for coreceptor binding and fusion. *Cell Host Microbe* 24, 579–592.
- Wang, H., Cohen, A.A., Galimidi, R.P., Gristick, H.B., Jensen, G.J., and Bjorkman, P.J. (2016). Cryo-EM structure of a CD4-bound open HIV-1 envelope trimer reveals structural rearrangements of the gp120 V1V2 loop. *Proc. Natl. Acad. Sci. USA* 113, E7151–E7158.
- Wang, H., Gristick, H.B., Scharf, L., West, A.P., Galimidi, R.P., Seaman, M.S., Freund, N.T., Nussenzweig, M.C., and Bjorkman, P.J. (2017). Asymmetric recognition of HIV-1 Envelope trimer by V1V2 loop-targeting antibodies. *Elife* 6, e27389.
- Ward, A.B., and Wilson, I.A. (2017). The HIV-1 envelope glycoprotein structure: nailing down a moving target. *Immunol. Rev.* 275, 21–32.
- Wibmer, C.K., Bhiman, J.N., Gray, E.S., Tumba, N., Abdool Karim, S.S., Williamson, C., Morris, L., and Moore, P.L. (2013). Viral escape from HIV-1 neutralizing antibodies drives increased plasma neutralization breadth through sequential recognition of multiple epitopes and immunotypes. *PLoS Pathog.* 9, e1003738.
- Wolbank, S., Kunert, R., Stiegler, G., and Katinger, H. (2003). Characterization of human class-switched polymeric (immunoglobulin M [IgM] and IgA) anti-human immunodeficiency virus type 1 antibodies 2F5 and 2G12. *J. Virol.* 77, 4095–4103.
- Yan, G.H., Wang, K., Shao, Z., Luo, L., Song, Z.M., Chen, J., Jin, R., Deng, X., Wang, H., Cao, Z., et al. (2018). Artificial antibody created by conformational reconstruction of the complementary-determining region on gold nanoparticles. *Proc. Natl. Acad. Sci. USA* 115, E34–E43.
- Zhu, P., Chertova, E., Bess, J., Jr., Lifson, J.D., Arthur, L.O., Liu, J., Taylor, K.A., and Roux, K.H. (2003). Electron tomography analysis of envelope glycoprotein trimers on HIV and simian immunodeficiency virus virions. *Proc. Natl. Acad. Sci. USA* 100, 15812–15817.
- Zhu, P., Liu, J., Bess, J., Jr., Chertova, E., Lifson, J.D., Grisé, H., Ofek, G.A., Taylor, K.A., and Roux, K.H. (2006). Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature* 441, 847–852.
- Zingler, K., and Littman, D.R. (1993). Truncation of the cytoplasmic domain of the simian immunodeficiency virus envelope glycoprotein increases env incorporation into particles and fusogenicity and infectivity. *J. Virol.* 67, 2824–2831.

## STAR★METHODS

### LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Rob Phillips ([phillips@pboc.caltech.edu](mailto:phillips@pboc.caltech.edu)).

#### Materials Availability Statement

This study did not generate new unique reagents.

### METHOD DETAILS

#### % Neutralization and HIV-1 Env Binding Probability

In this section, we determine the relationship between the experimentally measurable % neutralization and the probability  $p_{\text{bound}}$  (denoted by  $p_1 + p_2$  in the main text) that a trimeric Env spike will have a Fab bound to any of its three identical epitopes.

We begin by defining the typical % neutralization assay. We then analyze a linear model of HIV-1 infectivity discussed in the main text where the ability of the virus to infect a target cell is proportional to the number of Env spikes not bound by an antibody. We also discuss alternative models of infectivity, in which some minimal number of active spikes is required for a virion to infect a cell, finding that this model yields nearly identical predictions to the linear model.

Lastly, we investigate the importance of the experimentally-measured distribution of Env spikes on HIV. We start by assuming that each virion has the same number of spikes (given by the mean of the measured distribution) and then relax this assumption to characterize how infectivity changes when the number of Env spikes per virion is drawn from the full distribution.

#### Defining % Neutralization

In previous work, *in vitro* neutralization assays were carried out in 96 well plates, each well containing 250 TCID<sub>50</sub> and 25,000 cells that emit bioluminescence upon infection by the Pseudovirus as described (Montefiori, 2005). Antibodies and other potential inhibitors of neutralization are in vast excess over Pseudovirus and cells in these assays (e.g., 1 nM of antibody corresponds to 10<sup>11</sup> molecules/well). Upon infection, the cells emit light via a luciferase reporter as shown in Figure S1A. We define the percent of Pseudovirus neutralized as the fold-change in bioluminescence in the presence and absence of an inhibitor, namely,

$$\% \text{ neutralization} = 100 \frac{(\text{viral control} - \text{cell control}) - (\text{bioluminescence} - \text{cell control})}{\text{viral control} - \text{cell control}}, \quad (\text{Equation 10})$$

where *bioluminescence* is a measure of the light emitted in a well containing the pseudoviruses, cells, and antibodies; *viral control* is an assay using only cells and viruses (no antibodies) so that the cells emit maximal bioluminescence; and *cell control* is an assay using only cells (no viruses or antibodies) (Figure S1A).

The relative infectivity of a virion is defined as  $100 - (\% \text{ neutralization})$ , which equals 100% when no antibodies are present and 0% at saturating antibody concentrations where the binding sites on each HIV-1 spike are occupied.

From these assays, we define the equilibrium dissociation constant  $K_D^{(1)}$  characterizing the binding of the first diFab arm to HIV-1 Env (which is equivalent to the dissociation constant between a diFab with only one functional arm) as the midpoint of the neutralization curve. In this work, we chose the typical value  $K_D^{(1)} = 100$  nM from previous experiments.

#### A Linear Model of HIV-1 Infectivity

We now consider the linear model for HIV-1 infectivity used in the main text, which is predicated on the following assumptions: (i) each virus has the same number  $N = 14$  of Env trimers (or spikes), taken to be the mean of the experimentally measured distribution, (ii) each spike is active (able to help HIV-1 infect a target cell) if and only if none of its three subunits are bound by an antibody, and (iii) the relative infectivity of a virion is linearly proportional to its number of active spikes. Taken together, these assumptions imply that a virus with 7 active Env trimers shown on the left in Figure S1B will be half as infective as a completely unbound virus with 14 active trimers shown on the right.

The first assumption simplifies our analysis; below we relax this assumption and show that it minimally alters the % neutralization curves. The second assumption, that a spike is inactivated if at least one of its Fab sites is bound, is supported by experimental and computational studies (23). The third assumption has been observed when less than half the HIV-1 spikes are bound (27), but relative infectivity decreases faster than a linear model when more than half of the spikes are bound. In the following section, we relax this assumption and show that it minimally alters the % neutralization curves by making them slightly sharper. Taken together, these calculations demonstrate that our results are robust to the details of HIV-1 neutralization.

With these assumptions, a virion with  $n$  active spikes out of  $N = 14$  spikes total will have  $n/N$  the relative infectivity of a completely unbound virus, and hence the % neutralization is given by the expectation  $100(1 - n/N)$ , which equals 0% in the absence of antibodies when all spikes are active ( $n = N$ ) and equals 100% at saturating antibody concentrations when all spikes are inactive ( $n = 0$ ).

Given the probability  $p_{\text{bound}}$  that any spike will be bound, the probability of having  $n$  active (unbound) spikes equals  $p(n) = \binom{N}{n} (1 - p_{\text{bound}})^n p_{\text{bound}}^{N-n}$  and hence % neutralization is given by the weighted average of having  $n$  unbound spikes,

$$\begin{aligned}
 \% \text{ neutralization} &= 100 \sum_{n=0}^N p(n) \left(1 - \frac{n}{N}\right) \\
 &= 100 \sum_{n=0}^N \binom{N}{n} (1 - p_{\text{bound}})^n p_{\text{bound}}^{N-n} \left(1 - \frac{n}{N}\right) \\
 &= 100 - \frac{100}{N} \sum_{n=0}^N \binom{N}{n} (1 - p_{\text{bound}})^n p_{\text{bound}}^{N-n} n \tag{Equation 11} \\
 &= 100 - \frac{100}{N} N (1 - p_{\text{bound}}) \\
 &= 100 p_{\text{bound}},
 \end{aligned}$$

where in the fourth equality we used the average value of the binomial distribution. Therefore, the statistical mechanical model we develop to characterize the probability that a diFab is bound to Env also describes the ability of that diFab to neutralize a virion.

Equation 11 demonstrates that the exact number of spikes on an HIV-1 virion does not matter, implying that our framework will apply even in the face of the natural variation in spike density (Chertova et al., 2002; Zhu et al., 2003, 2006; Liu et al., 2008; Klein and Bjorkman, 2010). In particular, although the average distance between HIV-1 spikes is too large to inter-spike crosslink, the random positioning of these spikes on any virion may bring some subset of them sufficiently close to allow bivalent binding. Precluding the unlikely case where all spikes are precisely positioned to all be simultaneously crosslinked, Equation 11 guarantees that the remaining  $N$  spikes outside the subset (and hence viral neutralization) will conform to our model.

### Imposing a Hard Threshold for HIV-1 Infectivity

We now relax the third assumption stated above that relative infectivity is proportional to the number of active (i.e., unbound) HIV-1 spikes (Figure S1C, *Linear*). Instead, we posit that some minimum number of spikes must be active for a virion to be able to infect a target cell. This minimum number has been predicted to be between 1-3 active spikes (Brandenberg et al., 2017). Hence, we investigate two additional models where at least 3 of the HIV-1 spikes must be active for a virion to infect a target cell. In the first model (Figure S1C, *Linear with a Hard Threshold*), the relative infectivity increases linearly (and hence the % neutralization decreases linearly) with the number of active spikes  $>3$ , while in the second model (Figure S1C, *Hard Threshold*) we impose a pure threshold so that a virus is maximally infective provided at least 3 spikes are active.

For each model, we can alter the avidity factor  $\alpha$  to match the midpoints of each curve ( $\alpha = 1.5 \times 10^6$ ,  $10^6$ , or  $7 \times 10^6$  for the linear model, the linear model with a hard threshold, or the hard threshold model, respectively), since this parameter would ordinarily be inferred from a neutralization assay. As shown in Figure S1C, the linear model with a hard threshold is nearly identical to the linear model without this threshold, except that its % neutralization rises to 100% at lower antibody concentrations because it only needs to neutralize  $N - 2$  spikes to disable each virion. The hard-threshold model is sharper than the linear model, with the transition between no neutralization and full neutralization occurring when there are enough antibodies to bind  $N - 2$  spikes. Having noted these slight discrepancies, using any of these models would minimally affect our results, and hence we chose to proceed using the simplest linear response.

### % Neutralization Is Unchanged If the Number of env Spikes Varies between Virions

In this section, we relax the first assumption stated above and consider the number of spikes  $N$  on each virion to be drawn from a distribution ranging from 7-30 spikes per virion with an average of 14 (Chertova et al., 2002; Zhu et al., 2003, 2006; Liu et al., 2008; Klein and Bjorkman, 2010). We assume that the relative infectivity of a virus increases with each additional spike (with a maximum value attained by a virus with  $N_{\text{max}} = 30$  active spikes). The calculation for % neutralization follows analogously to Equation 11, except that the % neutralization of a virus with  $n$  active spikes is proportional to  $1 - n/N_{\text{max}}$  and that % neutralization must be averaged over all possible values of  $N$  drawn from its distribution, namely,

$$\begin{aligned}
 \% \text{ neutralization} &\propto \left\langle \sum_{n=0}^N \binom{N}{n} (1 - p_{\text{bound}})^n p_{\text{bound}}^{N-n} \left(1 - \frac{n}{N_{\text{max}}}\right) \right\rangle \\
 &= 1 - \frac{\langle N \rangle}{N_{\text{max}}} (1 - p_{\text{bound}}),
 \end{aligned} \tag{Equation 12}$$

In experimental measurements, % neutralization is always stretched to run from 0% to 100%; and if we apply this same stretching to Equation 12, we recover the result that % neutralization =  $100p_{\text{bound}}$  as in Equation 11.

This result is noteworthy in that it is independent of any details of the distribution (even the mean of the distribution does not affect neutralization when the function is stretched to run from 0% to 100%, as is customary practice in experimental measurements). However, this calculation rests upon the underlying assumption that relative infectivity increases linearly with the number of active spikes with no upper bound. If the relative infectivity saturates past a certain point (e.g., following a sigmoidal dependence), the result would no longer hold. Yet given that HIV-1 has so few spikes, it may turn out that each spike increases the relative infectivity of the virus by the same comparable amount.

### Number of Microstates in the Model including ssDNA

In this section, we enumerate the number of microstates associated with the three conformations of an HIV-1 Env spike – the unbound ( $\Omega_0$ ), singly bound ( $\Omega_1$ ), and bivalently bound states ( $\Omega_2$ ) – shown in [Figure S2](#). Our goal is to quantify the small fraction of linker configurations that enable a diFab to span the distance  $l_{\text{linker}}$  and bivalently bind.

The relative probability of each Env state is proportional to its Boltzmann weight,  $e^{-\beta(E_j - TS_j)} = \Omega_j e^{-\beta E_j}$ , where  $\beta = 1/k_B T$ ,  $E_j$  is the energy of state  $j$ ,  $T$  is the temperature (37°C),  $k_B$  = Boltzmann's constant,  $S = k_B \log \Omega_j$  is the entropy, and  $\Omega_j$  is the number of microstates of state  $j$ . Thus, the relative probabilities of the bivalent and monovalent states satisfy

$$\frac{\rho_2}{\rho_1} = \frac{\Omega_2}{\Omega_1} e^{-\beta(E_2 - E_1)} \quad (\text{Equation 13})$$

where  $\rho_j$ ,  $E_j$ , and  $\Omega_j$  represent the probability, energy, number of microstates when  $j$  Fabs are bound to an Env, respectively.

The number of configurations  $\Omega_1$  for a monovalently-bound diFab is proportional to the number of degrees of freedom available when one Fab is bound to its epitope and the second Fab is unbound but tethered to the first by DNA. More precisely, the different conformations arise from the rotation of the rigid dsDNA together with the small ssDNA motions. Since the associated entropy of this motion increases slowly with size,  $\Omega_1$  is effectively constant across all diFabs regardless of their linker composition. Similarly, since we assume that all diFabs are made up of the same Fab (e.g., all made using 3BNC117), the binding energy  $E_2 - E_1$  gained when the second Fab in a diFab binds to the HIV-1 spike is the same across all constructs. Thus, the only parameter in [Equation 13](#) that varies across diFabs with different linkers is the number of states  $\Omega_2$  of a bivalently bound linker, which we now proceed to compute.

### The Ideal Chain Model for ssDNA and dsDNA

We model the dsDNA as a 1D rigid rod and the ssDNA as a random walk with a step size given by its Kuhn length  $b_{\text{ssDNA}} = 2\xi_{\text{ssDNA}} = 3 \text{ nm} = 4.7$  bases ([Tinland et al., 1997](#); [Ambia-Garrido et al., 2010](#)). As in the main text,  $l_{\text{dsDNA}}$  denotes the total length of dsDNA while  $l_{\text{ssDNA}}$  denote the length of each of the two ssDNA segments.

Intuitively, because random walks tend to wander around their starting point, the optimal diFab will match its dsDNA length to the distance between the C-termini of two bound Fabs ( $l_{\text{dsDNA}} \approx l_{\text{linker}}$ ), to lose as little entropy as possibly when transitioning from a monovalently-bound to a bivalently-bound state. As shown in [Figure 2B](#) of the main text, a diFab that is too short ( $l_{\text{dsDNA}} \ll l_{\text{linker}}$ ) or too long ( $l_{\text{dsDNA}} \gg l_{\text{linker}}$ ) must stretch its ssDNA outwards or inwards to bivalently bind, thereby severely limiting the number of possible configurations in the doubly bound state. In the extreme limits where  $l_{\text{linker}} > l_{\text{dsDNA}} + 2l_{\text{ssDNA}} + l_{\text{flex}}$  or  $l_{\text{linker}} < l_{\text{dsDNA}} - 2l_{\text{ssDNA}} - l_{\text{flex}}$ , bivalent binding is impossible (recall that  $l_{\text{flex}}$  is the length by which the Fab can stretch as shown in [Figure 2A](#)). We will now make these statements precise by computing the probability that a linker configuration will permit a diFab to be bivalently bound.

### Computing the Probability of Bivalent Binding

We first turn our attention to the number of microstates of a bivalently bound linker ignoring the flexibility of the Fab shown in [Figure 3A](#) (i.e., in the  $l_{\text{flex}} = 0$  limit). Our goal will be to compute the probability  $p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, l_{\text{linker}})$  that the two ssDNA random walks and the dsDNA segment will span the appropriate distance  $l_{\text{linker}}$  necessary for the two Fabs to bivalently bind and then use this probability to count the number of microstates available for bivalent binding. In the last section, we consider the case where  $l_{\text{flex}} \neq 0$ .

As shown in [Figure S3A](#), each ssDNA random walk is composed of  $\hat{n} = l_{\text{ssDNA}}/b_{\text{ssDNA}}$  segments (where the notation distinguishes this variable from  $n$  in [Equation 2](#)) with Kuhn length  $b_{\text{ssDNA}} = 2\xi_{\text{ssDNA}}$ . The ssDNA and dsDNA in the linker must together span a fixed  $l_{\text{linker}}$ , where the direction and magnitude of this vector is determined by the geometry of the Env spike's epitopes. We now compute the probability that two ssDNA random walks sandwiched between a dsDNA rigid rod of size  $l_{\text{dsDNA}}$  spans  $l_{\text{linker}}$ . We proceed by considering four increasingly complex cases **Case 1**:  $l_{\text{dsDNA}} = 0$ ,  $l_{\text{linker}} = 0$ .

We begin by analyzing the special case of a diFab with an infinitesimally small dsDNA segment ( $l_{\text{dsDNA}} = 0$ ) binding to two epitopes that essentially lie on top of one another ( $l_{\text{linker}} = 0$ ). In other words, the constraint  $l_{\text{dsDNA}} = 0$  implies that both ssDNA random walks start off at the same location whereas  $l_{\text{linker}} = 0$  specifies that both random walks must end at the same location. This setup is shown in [Figure S3B](#) where the two random walks begin in the green square and end up within a small distance of each other represented by the gray cube.

Rather than analyzing the first random walk (with steps  $\vec{s}_1(1)$ ,  $\vec{s}_1(2)$ , and  $\vec{s}_1(3)$ ) and the second random walk (with steps  $\vec{s}_2(1)$ ,  $\vec{s}_2(2)$ , and  $\vec{s}_2(3)$ ) individually, we construct an effective random walk that traverses along one of the original walks and back along the other ( $\vec{s}_{\text{eff}}(j) = \vec{s}_1(j)$  for  $1 \leq j \leq 3$  and  $\vec{s}_{\text{eff}}(j) = -\vec{s}_2(j-3)$  for  $4 \leq j \leq 6$ ). This mapping is bijective, which means that every instance of the original random walks will correspond to a unique effective random walk and vice versa. Hence, the two original random walks will end at the same point if and only if the effective random walk ends near the origin. Therefore,  $p(0, l_{\text{ssDNA}}, 0)$  equals the probability that this effective random walk returns to the origin.

To make this argument precise, consider a 3D random walk starting at the origin and taking  $\hat{n}$  steps of length  $b_{\text{ssDNA}}$ . The probability that a random walk will end inside an infinitesimal volume  $dV$  centered at  $\vec{r}$  is given by (Berg, 1993)

$$P(\vec{r})dV = dV \left( \frac{3}{2\pi\hat{n}b_{\text{ssDNA}}^2} \right)^{\frac{3}{2}} e^{-\frac{3r^2}{2\hat{n}b_{\text{ssDNA}}^2}} \quad (\text{Equation 14})$$

Since each segment of the random walk can point in any direction, this probability only depends on the magnitude of  $\vec{r}$  and decreases exponentially with its distance from the origin. Note that  $P(\vec{r})$  is a probability density that upon multiplication by an infinitesimal volume  $dV$  denotes the probability of a random walk ending between  $\vec{r}$  and  $\vec{r} + d\vec{r}$ .

The probability that the effective random walk, formed by combining the two ssDNA random walks, will end up inside an infinitesimal volume  $dV$  around the origin is given by Equation 14 with  $\hat{n} \rightarrow 2\hat{n}$  and  $\vec{r} \rightarrow \vec{0}$ . Therefore, the fraction of bivalent binding configurations relative to monovalent binding configurations for two ssDNA random walks that start and end at the same location ( $\vec{l}_{\text{dsDNA}} = \vec{0}, \vec{l}_{\text{linker}} = \vec{0}$ ) is given by

$$p(0, l_{\text{ssDNA}}, 0) = P(\vec{0})dV = dV \left( \frac{3}{4\pi\hat{n}b_{\text{ssDNA}}^2} \right)^{\frac{3}{2}} \quad (\text{Equation 15})$$

Thus far, our calculation has been in terms of the probability  $P(\vec{0})dV$ . To convert this result into the number of microstates (and thereby compute the entropy), define  $\bar{\Omega}$  to be the number of microstates of each independent ssDNA segment (the  $2\hat{n}$  ssDNA segments are all assumed to rotate freely in the ideal chain model), the total number of microstates for the bivalent binding configurations is given by  $\bar{\Omega}^{2n} p(0, l_{\text{ssDNA}}, 0)$ . We note that this  $\bar{\Omega}$  will ultimately be subsumed into the  $\alpha$  factor in Equation 2 Case 2:  $\vec{l}_{\text{dsDNA}} = \vec{0}, \vec{l}_{\text{linker}} \neq \vec{0}$ .

We next consider the case where two ssDNA random chains with no interspersed dsDNA ( $\vec{l}_{\text{dsDNA}} = \vec{0}$ ) must end up at a displacement  $\vec{l}_{\text{linker}}$ . As shown in Figure S3C, the two ssDNA random walks are equivalent to the probability that a single random walk with  $2\hat{n}$  steps will finish at  $\vec{l}_{\text{linker}}$ . Therefore, the fraction of random walk configurations that allow bivalent binding when  $\vec{l}_{\text{dsDNA}} = \vec{0}$  and  $\vec{l}_{\text{linker}} \neq \vec{0}$  is given by Equation 14 with  $\hat{n} \rightarrow 2\hat{n}$  and  $\vec{r} \rightarrow \vec{l}_{\text{linker}}$ , namely,

$$p(0, l_{\text{ssDNA}}, l_{\text{linker}}) = P(\vec{l}_{\text{linker}})dV = dV \left( \frac{3}{4\pi\hat{n}b_{\text{ssDNA}}^2} \right)^{\frac{3}{2}} e^{-\frac{3l_{\text{linker}}^2}{4\hat{n}b_{\text{ssDNA}}^2}} \quad (\text{Equation 16})$$

As in Case 1, the total number of bivalent microstates is then given by  $\bar{\Omega}^{2n} p(0, l_{\text{ssDNA}}, l_{\text{linker}})$  Case 3:  $\vec{l}_{\text{dsDNA}} \neq \vec{0}, \vec{l}_{\text{linker}} = \vec{0}$ .

We next turn to a diFab whose two ssDNA ends must end at the same point ( $\vec{l}_{\text{linker}} = \vec{0}$ ) but whose linker contains dsDNA ( $\vec{l}_{\text{dsDNA}} \neq \vec{0}$ ) that can point in any direction. Analogous to the ssDNA segments, we assume the dsDNA has  $\bar{\Omega}$  microstates (discretized by solid angle so that the length of the dsDNA segment does not affect this number).

As shown in Figure S3D, the second random walk must end in a spherical shell of radius  $l_{\text{dsDNA}}$  surrounding the endpoint of the first random walk (note that we neglect the negligible width of the DNA double helix). We combine the two random walks into a single random walk with  $2\hat{n}$  steps by reversing the direction of the second random walk (pink) and translating it by  $-l_{\text{dsDNA}}$  so that it starts where the first random walk (purple) ends. Hence, the diFab can bind bivalently if and only if this effective random walk ends at a distance  $l_{\text{dsDNA}}$  from the origin.

Since  $P(\vec{r})$  is radially symmetric about the origin, the number of microstates for any particular orientation of the dsDNA will be  $\bar{\Omega}^{2\hat{n}} P(\vec{l}_{\text{dsDNA}})dV$  and the total number of microstates considering all dsDNA orientations will be  $\bar{\Omega}^{2\hat{n}+1} P(\vec{l}_{\text{dsDNA}})dV$ . However, in preparation for Case 4 below, it is instructive to compute the probability over linker states as the average over all dsDNA orientations,

$$\begin{aligned} p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, 0) &= \frac{dV}{4\pi l_{\text{dsDNA}}^2} \int_{\vec{x} \in B(\vec{0}, l_{\text{dsDNA}})} P(\vec{x}) d^2\vec{x} \\ &= P(l_{\text{dsDNA}})dV \\ &= dV \left( \frac{3}{4\pi\hat{n}b_{\text{ssDNA}}^2} \right)^{\frac{3}{2}} e^{-\frac{3l_{\text{dsDNA}}^2}{4\hat{n}b_{\text{ssDNA}}^2}} \end{aligned} \quad (\text{Equation 17})$$

where  $B(\vec{z}, r)$  represents the spherical shell of radius  $r$  centered at  $\vec{z}$ ,  $P(\vec{x})$  is given by Equation 14 with  $\hat{n} \rightarrow 2\hat{n}$ , and in the second step we used the radial symmetry of  $P(\vec{x})$ . The number of microstates is now given by  $\bar{\Omega}^{2\hat{n}+1} p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, 0)$  where the prefactor represents the  $\bar{\Omega}^{2\hat{n}}$  orientations of the ssDNA and the  $\bar{\Omega}$  orientations of the dsDNA. The similarity between Equations 16 and 17 reflect the symmetry between the dsDNA and the linker length in the system Case 4:  $\vec{l}_{\text{dsDNA}} \neq \vec{0}$ ,  $\vec{l}_{\text{linker}} \neq \vec{0}$ .

Finally, we turn to the case of a general diFab where the two ssDNA random walks are separated by a displacement  $\vec{l}_{\text{dsDNA}}$  of dsDNA and must end with displacement  $\vec{l}_{\text{linker}}$  from each other. As above, we transform these two random walks into a single effective random walk with  $2\hat{n}$  steps that must finish in a spherical shell centered at  $\vec{l}_{\text{linker}}$  with radius  $l_{\text{dsDNA}}$  (Figure S3E). Analogous to Case 3 above, the fraction of states of the dsDNA and ssDNA linker that allow bivalent binding is given by

$$p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, l_{\text{linker}}) = \frac{dV}{4\pi l_{\text{dsDNA}}^2} \int_{\vec{x} \in B(\vec{l}_{\text{linker}}, l_{\text{dsDNA}})} P(\vec{x}) d^2 \vec{x} \quad (\text{Equation 18})$$

where  $B(\vec{z}, r)$  represents the spherical shell of radius  $r$  centered at  $\vec{z}$  and  $P(\vec{x})$  is given by Equation 14 with  $\hat{n} \rightarrow 2\hat{n}$ . This final integral is straightforward to be evaluated analytically (see Data S1), yielding

$$p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, l_{\text{linker}}) = \frac{dV}{l_{\text{linker}} l_{\text{dsDNA}}} \sqrt{\frac{3}{16 \pi^3 \hat{n} b_{\text{ssDNA}}^2}} e^{-\frac{3(l_{\text{linker}}^2 + l_{\text{dsDNA}}^2)}{4\hat{n} b_{\text{ssDNA}}^2}} \sinh\left(\frac{3 l_{\text{linker}} l_{\text{dsDNA}}}{2 \hat{n} b_{\text{ssDNA}}^2}\right). \quad (\text{Equation 19})$$

This leads to Equation 2 with  $b_{\text{ssDNA}} = 2\xi_{\text{ssDNA}}$ ,  $\hat{n} = l_{\text{ssDNA}}/2\xi_{\text{ssDNA}}$ , and with the constant  $dV$  dropped (since it can be absorbed into  $\alpha$ ). Therefore, the number of configurations for the bivalently bound diFab linker is given by  $\Omega_2 = \bar{\Omega}^{2\hat{n}+1} p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, l_{\text{linker}})$ .

Note that  $p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, l_{\text{linker}})$  has no free parameters and is dictated purely by the geometry of each diFab. Furthermore, the factor  $p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, l_{\text{linker}})$  is the only term that varies between diFabs, whereas all remaining parameters (e.g.,  $K_D^{(1)}$ ,  $E_2 - E_1$ ) are the same across all constructs. Hence, it is the solely the loss of entropy contained in  $p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, l_{\text{linker}})$  that determines how much better one diFab will be than another. Lastly, we point out that Equation 19 is symmetric upon interchanging  $l_{\text{dsDNA}}$  and  $l_{\text{linker}}$ , since every state of the ssDNA in a bivalently bound diFab with  $l_{\text{dsDNA}}$  dsDNA and receptor binding sites spaced  $l_{\text{linker}}$  apart would also enabled bivalent binding (with the locations of the dsDNA and the receptor interchanged) of a diFab with a length  $l_{\text{linker}}$  of dsDNA binding to a receptor with binding sites spaced  $l_{\text{dsDNA}}$  apart.

### The Number of Bivalent versus Monovalent Microstates for the Linker

In this section, we now consider the flexibility  $l_{\text{flex}}$  of the diFab and compute the full expressions for the number of microstates  $\Omega_1$  and  $\Omega_2$  of the monovalently bound and bivalently-bound Env spike shown in Figure S3A. (We note that the number of microstates for the unbound state  $\Omega_0$  need not be computed explicitly because the ratio of entropy and energy between the unbound and monovalently bound states are quantified by  $K_D^{(1)}$  in Equation 1).

The number of microstates of the monovalently bound state is given by

$$\Omega_1 = \bar{\Omega}^{2\hat{n}+1}, \quad (\text{Equation 20})$$

where, as above,  $\bar{\Omega}$  denotes the microstates of each segment in the linker (the  $2\hat{n}$  ssDNA segments and the 1 dsDNA segment are all assumed to rotate freely in the ideal chain model). Note that this simple model neglects all interactions between the DNA, Fab, and Env including self-intersections.

When  $l_{\text{flex}} = 0$ , the multiplicity of the bivalently bound state was found above to be  $\bar{\Omega}^{2\hat{n}+1} p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, l_{\text{linker}})$  where  $l_{\text{linker}}$  represents the distance spanned by the linker. If we approximate the direction of flexibility of the Fabs and the line joining the Fabs' C-termini to be colinear, the number of bivalent microstates is given by

$$\begin{aligned} \Omega_2 &= \frac{\bar{\Omega}^{2\hat{n}+1}}{(\Delta l)^2} \int_{-\frac{l_{\text{flex}}}{2}}^{\frac{l_{\text{flex}}}{2}} \int_{-\frac{l_{\text{flex}}}{2}}^{\frac{l_{\text{flex}}}{2}} p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, l_{\text{linker}} + x_2 - x_1) dx_2 dx_1 \\ &\approx \frac{\bar{\Omega}^{2\hat{n}+1} l_{\text{flex}}^2}{(\Delta l)^2} p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, l_{\text{linker}}), \end{aligned} \quad (\text{Equation 21})$$

In the first equality,  $\Delta l$  represents a length scale that discretizes the flexibility imparted by the Fabs into the number of microstates. In the second equality, we assumed  $l_{\text{flex}} \ll l_{\text{linker}}$  (since  $l_{\text{flex}} \approx 20$  nm and  $l_{\text{linker}} \approx 1$  nm) so that  $l_{\text{linker}} + x_2 - x_1 \approx l_{\text{linker}}$  in the integrand. Substituting Equations 19 and 21 into Equation 13, the relative probability of the bivalent and monovalent states takes the form

$$\begin{aligned} \frac{p_2}{p_1} &= \frac{e^{-\beta(E_2 - E_1)} dV}{(\Delta l)^2 b_{\text{ssDNA}} l_{\text{linker}} l_{\text{dsDNA}}} \frac{l_{\text{flex}}^2}{l_{\text{dsDNA}}} \sqrt{\frac{3}{16 \pi^3 \hat{n}}} e^{-\frac{3(l_{\text{linker}}^2 + l_{\text{dsDNA}}^2)}{4\hat{n} b_{\text{ssDNA}}^2}} \sinh\left(\frac{3 l_{\text{linker}} l_{\text{dsDNA}}}{2 \hat{n} b_{\text{ssDNA}}^2}\right) \\ &\equiv \alpha n(l_{\text{dsDNA}}, l_{\text{ssDNA}}) \end{aligned} \quad (\text{Equation 22})$$

where we have defined the unitless prefactor  $\alpha = \frac{e^{-\beta(E_2 - E_1)} dV}{(\Delta l)^2 \rho_{\text{ssDNA}}}$  containing the constants that are independent of the diFab and Env and  $n(l_{\text{dsDNA}}, l_{\text{ssDNA}})$  to be the remaining geometry-dependent terms as per Equation 2. Using the normalization condition  $p_0 + p_1 + p_2 = 1$ , we find that the % neutralization of HIV-1 for a diFab linked together by dsDNA and ssDNA is given by

$$p_1 + p_2 = \frac{\frac{[\text{Ab}]}{K_D^{(1)}} + \frac{[\text{Ab}]}{K_D^{(1)}} \alpha n(l_{\text{dsDNA}}, l_{\text{ssDNA}})}{1 + \frac{[\text{Ab}]}{K_D^{(1)}} + \frac{[\text{Ab}]}{K_D^{(1)}} \alpha n(l_{\text{dsDNA}}, l_{\text{ssDNA}})}} \quad (\text{Equation 23})$$

As a final aside, we note that the ideal chain model is valid when there are at least 3 segments in the random walk (Figure S4), which requires  $3b_{\text{ssDNA}} \approx 14$  bases of ssDNA in the diFab. Constructs whose ssDNA has at least this many bases should be well characterized by our model, whereas diFabs with less ssDNA would require the more complicated worm-like chain model.

### Generalizing the Model to GappedFabs and TriFabs

In this section, we discuss how the model of diFab neutralization (Equation 4) can be generalized to account for either gappedFabs with ssDNA breaks within the dsDNA segment or triFabs that combine three Fabs to achieve greater avidity effects.

#### GappedFabs Represent a Random Walk of dsDNA

Here, we consider the effects of placing small ssDNA breaks within the dsDNA as shown in Figure S5A. For simplicity, we restrict our analysis to linkers composed of a middle segment with  $d_2$  bp dsDNA, surrounded by  $n_1$  segments of  $d_1$  bp dsDNA (each connected by short ssDNA segments;  $n_1 = 2$  in the three gappedFabs shown in Figure S5A), and flanked by  $s$  bases ssDNA on either side. This structure enables us to consider both the two outer ssDNA segments and the  $n_1$  dsDNA segments as random walks about the single dsDNA segment of length  $d_2$  (and it is straightforward to generalize to arbitrary gappedFab designs). In doing so, we assume the short ssDNA segments connecting the dsDNA act as free hinges with negligible lengths.

Equation 18 shows that in the case  $n_1 = 0$  where the linker must span the distance  $l_{\text{linker}}$ , the combined ssDNA random walk ( $2n$  segments of length  $b_{\text{ssDNA}}$ ) starting at the origin must end on a sphere of radius  $l_{\text{dsDNA}}$  around  $l_{\text{linker}}$ . When  $n_1 > 0$ , the combined ssDNA random walk ends at the arbitrary point  $\vec{z}$ , and the dsDNA random walk must then start at  $\vec{z}$  and end on a sphere of radius  $l_{d_2} = d_2(0.34 \text{ nm / bp})$  around  $l_{\text{linker}}$ , representing the length of the  $d_2$  segment. Using the probability density Equation 14 for a random walk, the probability that the linker in a gappedFab will bivalently bind is given by

$$p_{\text{gap}}(l_{\text{dsDNA}}, l_{\text{ssDNA}}, l_{\text{linker}}) = \int_{\vec{x} \in B(\vec{l}_{\text{linker}}, l_{d_2})} \int_{\vec{z} \in \mathbb{R}^3} P_{\text{ssDNA}}(\vec{z}) P_{\text{dsDNA}}(\vec{x} - \vec{z}) d^3 \vec{z} d^3 \vec{x} \quad (\text{Equation 24})$$

where

$$P_{\text{ssDNA}}(\vec{z}) = \left( \frac{3}{2\pi b_{\text{ssDNA}}^2 (2n)} \right)^{\frac{3}{2}} e^{-\frac{3z^2}{2b_{\text{ssDNA}}^2 (2n)}} \quad (\text{Equation 25})$$

represents the probability that the endpoint of the ssDNA random walk starting at the origin will end at  $\vec{z}$ ,

$$P_{\text{dsDNA}}(\vec{x} - \vec{z}) = \left( \frac{3}{2\pi l_{d_1}^2 n_1} \right)^{\frac{3}{2}} e^{-\frac{3|\vec{x} - \vec{z}|^2}{2l_{d_1}^2 n_1}} \quad (\text{Equation 26})$$

denotes the probability that the endpoint of the dsDNA random walk starting at  $\vec{z}$  will end at  $\vec{x}$ , and  $l_{d_1} = d_1(0.34 \text{ nm / bp})$ . Replacing  $p \rightarrow p_{\text{gap}}$  and  $\bar{\Omega}^{2n} \rightarrow \bar{\Omega}^{2n+n_1}$  in Equation 21 yields the desired number of microstates for the bivalent configuration.

Figure S6 shows how the potency of the previously constructed diFab ( $d=62, s=12$ ) gets worse (larger  $\text{IC}_{50}$ ) as the dsDNA segment is broken up into  $\hat{n}$  pieces of equal length ( $n_1 = \hat{n} - 1, d_1 = d_2 = \frac{62}{\hat{n}}, s = 12$ ). Since the  $\hat{n}$  segments of dsDNA will be much more confined when the diFab is bivalently bound relative to the monovalent configuration, the entropic penalty of bivalent binding quickly increases with the number of segments leading to lower potency.

#### TriFabs Exhibit Greater Avidity Than diFabs

Previously, we showed that the diFab with ( $d=62, s=12$ ) has an  $\text{IC}_{50}$  that is 100-fold smaller than its one-armed Fab counterpart (Galimidi et al., 2015). This raises the question of whether a construct with additional Fabs could further reduce the  $\text{IC}_{50}$ . Here, we outline how our model can be extended to consider a linear triFab (shown in Figure 1B), and this calculation can be readily extended to the alternate triFab designs in Figure S5B. The calculation below predicts that the optimal triFab will be 100-fold more potent than the optimal diFab, providing a method to leverage the knowledge of the HIV-1 Envelope spike we derived from our synthetic diFabs to engineer even more potent reagents.

We model the ssDNA segments with  $s$  bases as random walks and neglect both the self-intersection of these random walks as well as intersections with the Env spike, though we note that these effects may be more prominent in a triFab than in a diFab. Furthermore, we neglect the combinatorics characterizing which binding arm attaches to an Env binding site (e.g., there are  $\binom{3}{1}$  ways to

bind monovalently;  $\binom{3}{2}$  ways to bind bivalently, although in the linear triFab the simultaneous binding of nearest neighbor arms will be different from the binding of the two arms furthest apart; and  $\binom{3}{3}$  ways to bind trivalently) as well as the details of these configurations (e.g., there are six possible configurations of binding the three triFab arms to the three Env epitopes).

With these assumptions, the Boltzmann statistical weights for the triFab (analogous to those in Figure 2A for the diFab) are  $1$ ,  $\frac{[\text{Ab}]}{K_D^{(1)}}$ ,  $\frac{[\text{Ab}]}{K_D^{(1)}} \alpha n$ , and  $\frac{[\text{Ab}]}{K_D^{(1)}} \alpha^2 n^2$  for the states with 0, 1, 2, and 3 Fab arms bound, respectively. Analogous to Equation 23, the probability of binding (and hence neutralizing) a spike is given by

$$p_1 + p_2 + p_3 = \frac{\frac{[\text{Ab}]}{K_D^{(1)}} + \frac{[\text{Ab}]}{K_D^{(1)}} \alpha n + \frac{[\text{Ab}]}{K_D^{(1)}} \alpha^2 n^2}{1 + \frac{[\text{Ab}]}{K_D^{(1)}} + \frac{[\text{Ab}]}{K_D^{(1)}} \alpha n + \frac{[\text{Ab}]}{K_D^{(1)}} \alpha^2 n^2}} \quad (\text{Equation 27})$$

with an  $\text{IC}_{50}$  given by

$$\text{IC}_{50}^{\text{triFab}} = \frac{K_D^{(1)}}{1 + \alpha n + \alpha^2 n^2}. \quad (\text{Equation 28})$$

As described in the main text, this implies that the boost in avidity in going from an optimal diFab to an optimal triFab is equal to the boost in going from a Fab to an optimal diFab.

#### DATA AND CODE AVAILABILITY

The Supplemental Information includes six figures as well as a Mathematica notebook that reproduces the calculations and plots in this manuscript. All files can be found in the online version of this article.