Supporting Information

A. % Neutralization and the Probability Envelope Protein is Bound

In this section, we determine the relationship between the experimentally measured % neutralization assay and the probability p_{bound} (denoted by $p_1 + p_2$ in the main text) that an Env trimer spike will have an antibody bound to any of its three identical CD4 binding sites.

We begin by defining how % neutralization is measured experimentally. We then analyze a linear model of HIV-1 infectivity discussed in the main text where the ability of HIV-1 to infect a target cell is proportional to the number of Env spikes not bound by antibodies. We also discuss alternate models of infectivity where some minimal number of active spikes are required for a virion to infect a cell and find 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 begin with a simple model where each virion has the mean number of spikes dictated by this distribution. We then characterize the assumptions under which the resulting infectivity of the virus will not change if the number of spikes per virion is drawn from the full distribution.

Defining % Neutralization
Pseudovirus preparations are titered to determine the tissue culture infectious dose (TCID$_{50}$). In vitro neutralization assays were done in 96 well plates, each well containing 250 TCID$_{50}$ and 25,000 cells that emit bioluminescence upon infection by the pseudovirus as described (40). Antibodies and other potential inhibitors of neutralization are in vast excess over pseudovirus and cells in these assays (e.g., 1 nM antibody corresponds to 10^{11} molecules/well). Upon infection, the cells emit light via a luciferase reporter as shown in Fig. S1A. We define the percent of pseudovirus neutralized as the fold-change in bioluminescence in the presence and absence of an inhibitor, namely,

$$\% \text{ neutralization} = 100 \frac{(\text{viral control} - \text{cell control}) - (\text{bioluminescence} - \text{cell control})}{\text{viral control} - \text{cell control}}, \quad [\text{S1}]$$

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) (Fig. 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 by an inhibitor.

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 none of its three Fab binding sites are occupied by an antibody and inactive otherwise, 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 seven active Env trimers shown on the left in Fig. 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 does not alter 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 may leads to slightly sharper % neutralization curves that agree less well with the data.

Assuming the assumptions outlined above, a virion with n active spikes out of $N = 14$ spikes total will have $\frac{n}{N}$ the relative infectivity of a completely unbound virus, and hence the % neutralization is given by the expectation $100 \langle 1 - \frac{n}{N} \rangle$ 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_{bound} that any spike will be bound, the probability of having n active (unbound) spikes equals $\binom{N}{n}(1 - p_{\text{bound}})^n p_{\text{bound}}^{N-n}$ and hence % neutralization is given by

$$
\begin{aligned}
\% \text{ neutralization} &= 100 \sum_{n=0}^{N} \binom{N}{n}(1 - p_{\text{bound}})^n p_{\text{bound}}^{N-n} \left(1 - \frac{n}{N}\right) \\
&= 100 - 100 \sum_{n=0}^{N} \binom{N}{n}(1 - p_{\text{bound}})^n p_{\text{bound}}^{N-n} \frac{n}{N} \\
&= 100 - 100(1 - p_{\text{bound}}) \sum_{m=0}^{N-1} \binom{N-1}{m}(1 - p_{\text{bound}})^m p_{\text{bound}}^{N-1-m} \\
&= 100\, p_{\text{bound}}, \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\text{[S2]}
\end{aligned}
$$

where in the third equality we changed variables to $m = n - 1$. Therefore, we see that the formula Eq. 9 in the main text quantifying the probability that any of the Fab epitopes on an HIV-1 trimer will be bound allows us to characterize the neutralization data in Fig. 2D.
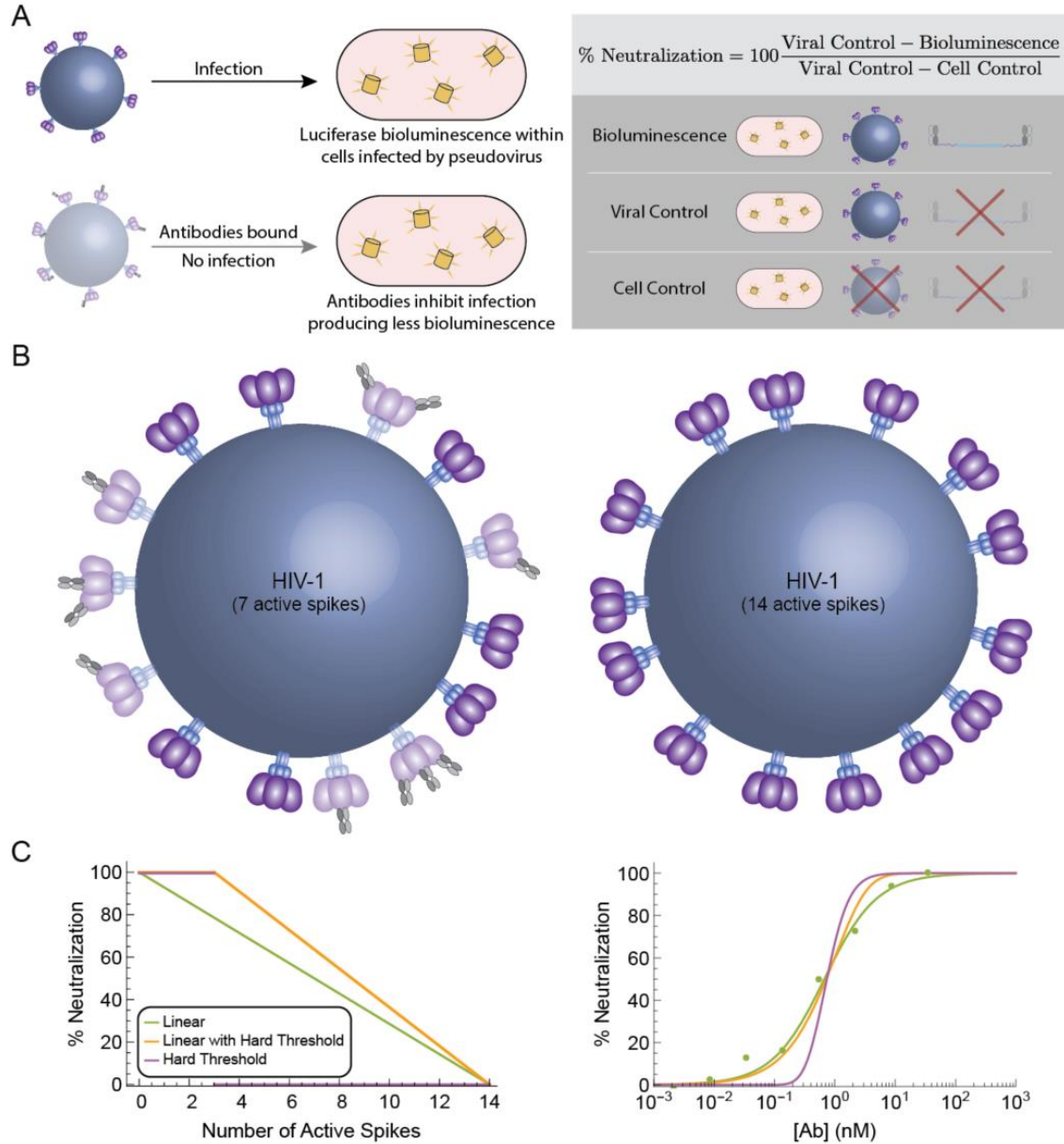
Figure S1. A Linear model of HIV-1 neutralization. (A) Neutralization of HIV-1 pseudovirus was assayed by evaluating reporter cells that emit light upon HIV-1 infection via luciferase (1). Adding an inhibitor (e.g., a diFab) results in decreased bioluminescence. (B) We model each HIV-1 virion as having 14 spikes that are inactivated (represented as partially transparent spikes) when a 3BNC60 Fab is bound to any of its three binding sites on an Env trimer. In the main text, we assume a linear model in which the infectivity of a virion is proportional to its number of active spikes; for example, the virus on the right will be twice as infective as the virus on the left. (C) Different models for % neutralization (or the relative infectivity given by $100 - \%$ neutralization) as a function of the number of active (unbound) HIV-1 spikes (left) and their corresponding % neutralization curves (right). The geometric factor $\tilde{\alpha}$ in Eq. 8 quantifying the effects of diFab avidity was adjusted for each model (5×10^5 for the linear model; 3×10^5 for the linear model with a hard threshold; 20×10^5 for the hard threshold model) to match the data. The values of the remaining parameters were the same as in Fig. 2.

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 (Fig. 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 (49). 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 (Fig. 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 (Fig. 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 $\tilde{\alpha}$ to best match the data ($\tilde{\alpha} = 5 \times 10^5$, 3×10^5, or 20×10^5 for the linear model, the linear model with a hard threshold, or the hard threshold model, respectively), as shown in Fig. S1C for the (d=62, s=12) diFab. As expected, 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. While we note that data from in vitro neutralization assays are inherently noisy, the shallower linear response characterizes the data slightly better. Lastly we note that the neutralization profiles of other diFabs would have the same shape as the (d=62, s=12) curve and that the spacing between the other diFab neutralization curves is unchanged by the relationship between % neutralization and the number of active spikes. Therefore, using any of these models would minimally affect our results.

% 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 (22-26). 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 Eq. S2, except that the % neutralization of a virus with n active spikes is proportional to $1 - \frac{n}{N_{\text{max}}}$ and that % neutralization must be averaged over all possible values of N drawn from its distribution, namely,

$$\% \text{ neutralization} \propto \langle \sum_{n=0}^{N} \binom{N}{n} (1 - p_{\text{bound}})^n p_{\text{bound}}^{N-n} \left(1 - \frac{n}{N_{\text{max}}}\right) \rangle$$

$$= 1 - \frac{\langle N \rangle}{N_{\text{max}}} (1 - p_{\text{bound}}). \tag{S3}$$

However, this function does not run from 0% to 100% (as does % neutralization), and upon stretching this function to span these two limits we recover the result that $\% \text{ neutralization} = 100 p_{\text{bound}}$ as in Eq. S2.

This result is noteworthy in that it is independent of any details of the distribution (even the mean gets cancelled when the function is stretched to run from 0% to 100%). However, it 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. However, 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.

B. Number of Microstates in a Model Excluding ssDNA

In this section, we count the number of microstates available for bivalent binding in the simple model in which contributions from the ssDNA segments are ignored within the linker (Fig. 2A). The flexibility of the Fabs between the C_H1-C_L and V_H-V_L domains allows it to bend by a distance l_{flex}. To simplify the calculation, we will only consider the motion of l_{flex} in the direction colinear with the dsDNA strand so that the problem becomes one dimensional, with the relevant lengths shown in Fig. S2A. We define $n(l_{dsDNA})$ as the length along which the dsDNA can translate while the diFab is bivalently bound. Analyzing the system using a 1D lattice model where Δl is the discretized length scale, the number of microstates for the bivalently bound configuration will be $\frac{n(l_{dsDNA})}{\Delta l}$.

When the dsDNA is too long ($l_{dsDNA} \geq l_{linker} + l_{flex}$) or too short ($l_{dsDNA} \leq l_{linker} - l_{flex}$), the diFab cannot bivalently bind and hence there are zero microstates ($n(l_{dsDNA}) = 0$) in these limits. In between these two limits, bivalent binding is possible.

Fig. S2B shows the case when the dsDNA is longer than the linker length ($l_{dsDNA} \geq l_{linker}$). When bivalently bound, the left-most position of the dsDNA occurs when the left Fab is flexing to its left-most position; the right-most position of the dsDNA similarly occurs when the right Fab is flexing to its right-most position. Therefore, the bivalently bound state will allow the dsDNA to translate by a distance $n(l_{dsDNA}) = l_{flex} - (l_{dsDNA} - l_{linker})$.

Fig. S2C shows the corresponding case when the dsDNA is shorter than the linker ($l_{dsDNA} \leq l_{linker}$). In this case, the left-most position of the dsDNA is dictated by the right Fab stretching to its left-most position, with a similar statement holding for the right-most DNA position. Hence, the dsDNA can slide by however much longer it is than the minimum distance $l_{linker} - l_{flex}$ that allows bivalent binding, namely, $l_{dsDNA} - (l_{linker} - l_{flex})$. Putting the two cases together, the general expression or the length that the dsDNA can translate within the bivalent binding configurations is given by

$$n(l_{dsDNA}) = \max(l_{flex} - |l_{dsDNA} - l_{linker}|, 0), \qquad [S4]$$

leading to the number of microstates in Eq. 3, $\Omega_2 = \frac{n(l_{dsDNA})}{\Delta l}$.

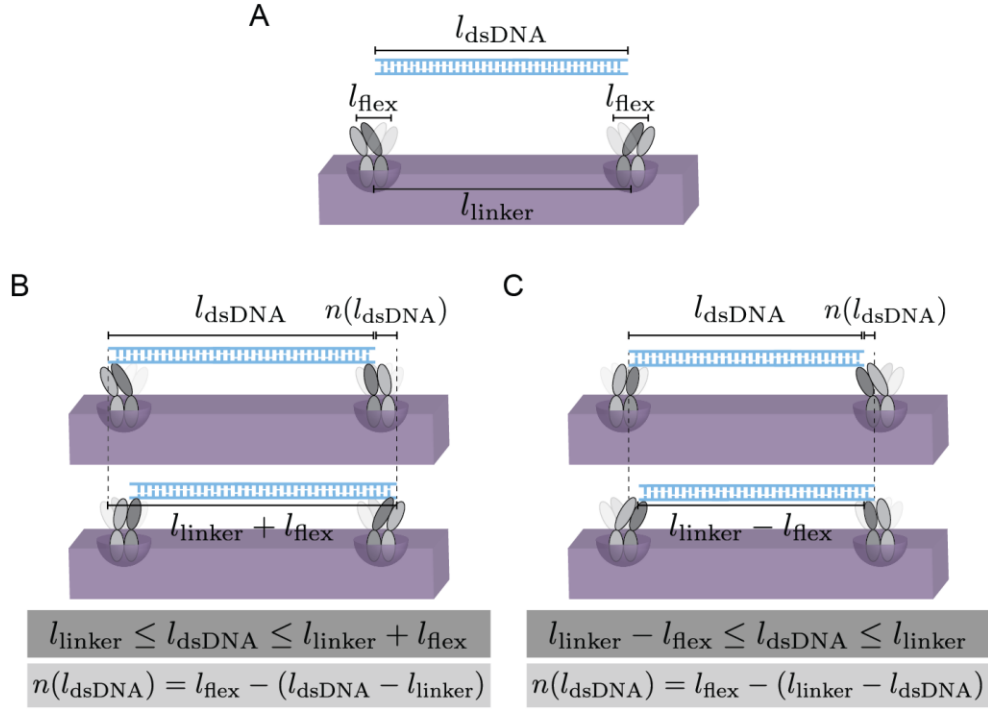**Figure S2. Number of bivalent microstates in the model excluding ssDNA.** (A) dsDNA of length l_{dsDNA} must bridge the distance l_{linker} between the C-termini of two bound Fabs. The C_H1-C_L and V_H-V_L hinge between the two Fabs is flexible, permitting the C-terminal C_H1 domain residue to which the DNA is attached to flex about its central point by a distance l_{flex}. (B) When the dsDNA is longer than the linker length, the left Fab can flex to its left-most point and the dsDNA can translate to the right until the right-most Fab reaches its right-most point. The length of the translation permitted when a diFab is bivalently bound is denoted by $n(l_{\mathrm{dsDNA}})$. (C) When the dsDNA is shorter than the linker length, the left-most position of the dsDNA is dictated by the right Fab stretching to its left-most point; the dsDNA can translate to the right until the left Fab flexes to its right-most point.

C. Number of Microstates in the Model Including ssDNA

In this section, we enumerate the number of microstates associated with the three states of an HIV-1 Env spike – the unbound state (Ω_0), the singly bound state (Ω_1), and the bivalently bound state (Ω_2) – shown in Fig. S3A for the model where we include both the dsDNA and the ssDNA in the diFab linker.
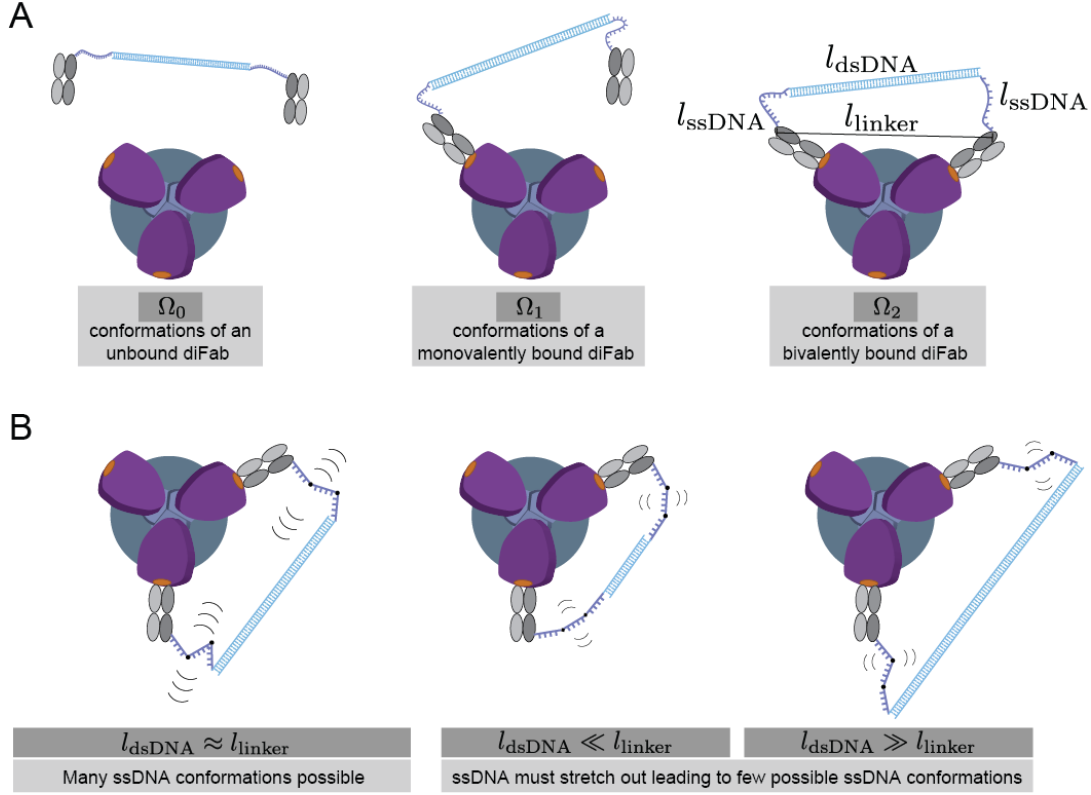

Figure S3. Conformational states of a diFab. (A) To bivalently bind, the linker region of a diFab (a length l_{dsDNA} of dsDNA flanked by two lengths l_{ssDNA} of ssDNA) must span a distance l_{linker} that is dictated by the Env protein's structure. The dsDNA is a rigid rod while the ssDNA (shown floppy in Panel A) is modeled as an ideal chain (shown as straight line segments in Panel B). (B) The optimal tether will match $l_{\text{dsDNA}} \approx l_{\text{linker}}$, because when $l_{\text{dsDNA}} \ll l_{\text{linker}}$ or $l_{\text{dsDNA}} \gg l_{\text{linker}}$ the ssDNA must be extended, severely limiting the number of conformational states available and inducing a much larger entropic penalty cost for bivalent binding.

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 $b_{\text{ssDNA}} = 2\xi_{\text{ssDNA}}$. While the Kuhn length of free ssDNA is $b_{\text{ssDNA}} = 3 \text{ nm} = 4.7 \text{ bases}$ (43, 44), we will determine the Kuhn length in our system through nonlinear fitting. As in the main text, l_{dsDNA} and l_{ssDNA} denote the lengths of the dsDNA and ssDNA, respectively.

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 terminals 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 Fig. S3B, 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 $l_{\text{linker}} > l_{\text{dsDNA}} + 2l_{\text{ssDNA}}$ or $l_{\text{linker}} < l_{\text{dsDNA}} - 2l_{\text{ssDNA}}$, bivalent binding is impossible. 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 Fig. 2 (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_{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 of this Appendix, we consider the case of $l_{\text{flex}} \neq 0$.

As shown in Fig. S4A, each ssDNA random walk is composed of $n = \frac{l_{\text{ssDNA}}}{b_{\text{ssDNA}}}$ segments with Kuhn length $b_{\text{ssDNA}} = 2\xi_{\text{ssDNA}}$ given by twice the persistence length. The ssDNA and dsDNA in the linker must together span a fixed $\vec{l}_{\text{linker}}$, 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 $\vec{l}_{\text{dsDNA}}$ spans $\vec{l}_{\text{linker}}$. We proceed by considering four increasingly complex cases.

<u>Case 1:</u> $\vec{l}_{\text{dsDNA}} = \vec{0}, \vec{l}_{\text{linker}} = \vec{0}$
We begin by analyzing the special case of a diFab with an infinitesimally small dsDNA segment ($\vec{l}_{\text{dsDNA}} = \vec{0}$) binding to two epitopes that essentially lie on top of one another ($\vec{l}_{\text{linker}} = \vec{0}$). In other words, the constraint $\vec{l}_{\text{dsDNA}} = \vec{0}$ implies that both ssDNA random walks start off at the same location whereas $\vec{l}_{\text{linker}} = \vec{0}$ specifies that both random walks must end at the same location. This setup is shown in Fig. S4B 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. This implies that 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 n steps of length b_{ss}. The probability that a random walk will end inside an infinitesimal volume dV centered at $\vec{r}$ is given by (77)

$$P(\vec{r})dV = dV \left(\frac{3}{2\pi n b_{\text{ssDNA}}^2}\right)^{\frac{3}{2}} e^{-\frac{3r^2}{2n b_{\text{ssDNA}}^2}}. \qquad [S5]$$

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 $\vec{r} + d\vec{r}$.
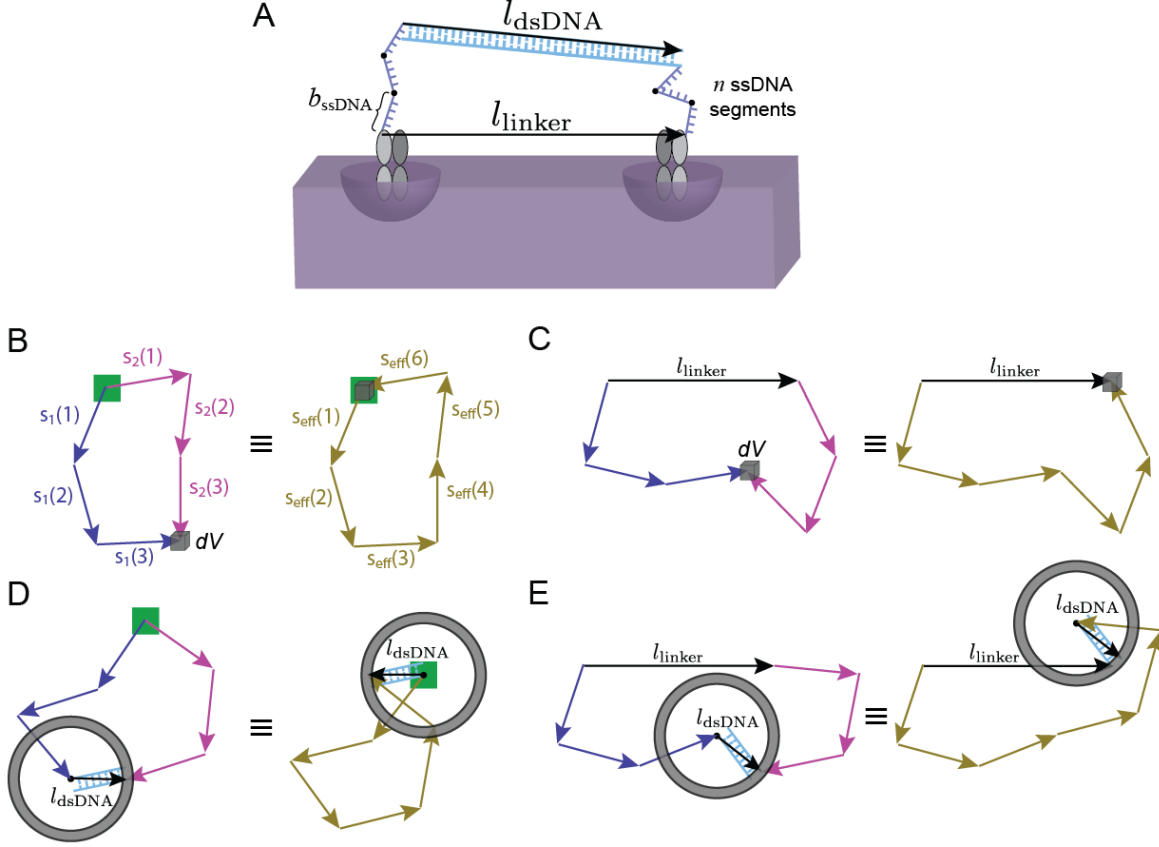
Figure S4. The ideal chain model for a diFab linker. (A) We compute the probability $p(l_{dsDNA}, l_{ssDNA}, l_{linker})$ that the ssDNA (n segments of length b_{ssDNA}) and dsDNA (1 rigid segment of length l_{dsDNA}) in the diFab linker will end at the appropriate distance $\vec{l}_{linker}$ required for bivalent binding. We consider the following cases: (B) $\vec{l}_{dsDNA} = \vec{0}, \vec{l}_{linker} = \vec{0}$; (C) $\vec{l}_{dsDNA} = \vec{0}, \vec{l}_{linker} \neq \vec{0}$; (D) $\vec{l}_{dsDNA} \neq \vec{0}, \vec{l}_{linker} = \vec{0}$; and the full model (E) $\vec{l}_{dsDNA} \neq \vec{0}, \vec{l}_{linker} \neq \vec{0}$. The two random walks start at each end of $\vec{l}_{linker}$, which is denoted by a green square when $\vec{l}_{linker} = \vec{0}$ (Panels B and D). Random walks end in the same location if they are within a small volume dV of each other (Panels B and C) or if their ends are within dV of the dsDNA of length l_{dsDNA} (Panels D and E). In each case, the two ssDNA random walks are combined into a single effective random walk with $2n$ steps.

The probability that the effective random walk, formed by combining the two ssDNA random walks, will end up inside an infinitesimal volume dV from the origin is given by Eq. S5 with $n \rightarrow 2n$ 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 ($\vec{l}_{dsDNA} = \vec{0}, \vec{l}_{linker} = \vec{0}$) is given by

$$p(0, l_{ssDNA}, 0) = P(\vec{0})dV = dV \left(\frac{3}{4\pi n b_{ssDNA}^2} \right)^{\frac{3}{2}}. \qquad [S6]$$

Defining $\bar{\Omega}$ to be the number of microstates of each independent ssDNA segment (the $2n$ 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 $\bar{\Omega}^{2n} p(0, l_{ssDNA}, 0)$.

<u>Case 2:</u> $\vec{l}_{\text{dsDNA}} = \vec{0}, \vec{l}_{\text{linker}} \neq \vec{0}$

We next consider the case where two ssDNA random chains with no interspersed dsDNA ($\vec{l}_{\text{dsDNA}} = 0$) must end up at a displacement $\vec{l}_{\text{linker}}$. As shown in Fig. S4C, the two ssDNA random walks are equivalent to the probability that a single random walk with $2n$ 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 Eq. S5 with $n \to 2n$ and $\vec{r} \to \vec{l}_{\text{linker}}$, namely,

$$p(0, l_{\text{ssDNA}}, l_{\text{linker}}) = P(\vec{l}_{\text{linker}})dV = dV \left(\frac{3}{4\pi n b_{\text{ssDNA}}^2}\right)^{\frac{3}{2}} e^{-\frac{3l_{\text{linker}}^2}{4nb_{\text{ssDNA}}^2}}. \qquad [S7]$$

As in case 1, the total number of bivalent microstates is then given by $\bar{\Omega}^{2n}p(0, l_{\text{ssDNA}}, l_{\text{linker}})$.

<u>Case 3</u>: $\vec{l}_{\text{dsDNA}} \neq \vec{0}, \vec{l}_{\text{linker}} = \vec{0}$

We next turn to a diFab whose two ssDNA ends must end at the same point ($l_{\text{linker}} = 0$) 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 Fig. S4D, the second random walk must end in a spherical shell of radius l_{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 $2n$ steps by reversing the direction of the second random walk (pink) and translating it by $-\vec{l}_{\text{dsDNA}}$ 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_{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 $\bar{\Omega}^{2n}P(\vec{l}_{\text{dsDNA}})dV$ and the total number of microstates considering all dsDNA orientations will be $\bar{\Omega}^{2n+1}P(\vec{l}_{\text{dsDNA}})dV$. However, in preparation for case 4 below, it is instructive to compute the probability over linker states as the average over all dsDNA orientations,

$$
\begin{aligned}
p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, 0) &= \frac{dV}{4\pi l_{\text{dsDNA}}^2} \int_{\vec{x} \in B(0, l_{\text{dsDNA}})} P(\vec{x}) d^2\vec{x} \\
&= P(l_{\text{dsDNA}})dV \\
&= dV \left(\frac{3}{4\pi n b_{\text{ssDNA}}^2}\right)^{\frac{3}{2}} e^{-\frac{3l_{\text{dsDNA}}^2}{4nb_{\text{ssDNA}}^2}}, \qquad [S8]
\end{aligned}
$$

where $B(\vec{z}, r)$ represents the spherical shell of radius r centered at $\vec{z}$, $P(\vec{x})$ is given by Eq. S5 with $n \to 2n$, and in the second step we used the radial symmetry of $P(\vec{x})$. The number of microstates is now given by $\bar{\Omega}^{2n+1}p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, 0)$ where the prefactor represents the $\bar{\Omega}^{2n}$ orientations of the ssDNA and the $\bar{\Omega}$ orientations of the dsDNA. The similarity between Eqs. S7 and S8 reflect the symmetry between the dsDNA and the linker length in the system.

<u>Case 4:</u> $\vec{l}_{\text{dsDNA}} \neq \vec{0}, \vec{l}_{\text{linker}} \neq \vec{0}$

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 $2n$ steps that must finish in a spherical shell centered at $\vec{l}_{\text{linker}}$ with radius l_{dsDNA} (Fig. S4E).

Analogous to case 3 above, the fraction of states of the dsDNA and ssDNA linker that allow bivalent binding is given by

$$p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, l_{\text{linker}}) = \frac{dV}{4\pi\, l_{\text{dsDNA}}^2} \int_{\vec{x} \in B(\vec{l}_{\text{linker}}, r)} P(\vec{x})\, d^2\vec{x} \qquad [S9]$$

where $B(\vec{z}, r)$ represents the spherical shell of radius r centered at $\vec{z}$ and $P(\vec{x})$ is given by Eq. S5 with $n \to 2n$. This final integral is straightforward to evaluated analytically (see the Supplementary Mathematica notebook), yielding

$$p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, l_{\text{linker}}) = \frac{dV}{l_{\text{linker}}\, l_{\text{dsDNA}}} \sqrt{\frac{3}{16\pi^3 n b_{\text{ssDNA}}^2}}\, e^{-\frac{3(l_{\text{linker}}^2 + l_{\text{dsDNA}}^2)}{4 n b_{\text{ssDNA}}^2}} \sinh\left(\frac{3\, l_{\text{linker}}\, l_{\text{dsDNA}}}{2 n b_{\text{ssDNA}}^2}\right). \quad [S10]$$

This leads to Eq. 6 with $b_{\text{ssDNA}} = 2\xi_{\text{ssDNA}}$, $n = \frac{l_{\text{ssDNA}}}{2\xi_{\text{ssDNA}}}$, and with the constant dV dropped (since it will be absorbed into $\tilde{\alpha}$). Therefore, the number of configurations for the bivalently bound diFab linker is given by $\Omega_2 = \overline{\Omega}^{2n+1} p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, l_{\text{linker}})$.

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 Eq. S10 is symmetric upon interchanging l_{dsDNA} and l_{linker}, since every state of the ssDNA in a bivalently bound diFab with l_{dsDNA} dsDNA and receptor binding sites spaced l_{linker} apart would also enabled bivalent binding (with the locations of the dsDNA and the receptor interchanged) of a diFab with a length l_{linker} of dsDNA binding to a receptor with binding sites spaced l_{dsDNA} apart.

The Number of Bivalent versus Monovalent Microstates for the Linker
In this section, we compute the full expressions for the number of microstates Ω_1 and Ω_2 of the monovalently bound and bivalently-bound Env spike shown in Fig. S3A (including the flexibility l_{flex}), which will permit us to compute the relative probability of these two states using Eq. 2. We note that the number of microstates for the unbound state Ω_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 Eq. 1.

The number of microstates of the monovalently bound state is given by
$$\Omega_1 = \overline{\Omega}^{2n+1}, \qquad [S11]$$
where, as above, $\overline{\Omega}$ denotes the microstates of each segment in the linker (the $2n$ 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 $\overline{\Omega}^{2n+1} p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, l_{\text{linker}})$ where l_{linker} represents the distance spanned by the linker. Making the same approximation as in Appendix B that the direction of flexibility of the Fabs and the line joining the Fabs' C-termini are colinear, the number of bivalent microstates is given by

$$\Omega_2 = \frac{\overline{\Omega}^{2n+1}}{(\Delta l)^2} \int_{-\frac{l_{\text{flex}}}{2}}^{\frac{l_{\text{flex}}}{2}} \int_{-\frac{l_{\text{flex}}}{2}}^{\frac{l_{\text{flex}}}{2}} p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, l_{\text{linker}} + x_2 - x_1)\, dx_2\, dx_1$$

$$\approx \frac{\overline{\Omega}^{2n+1} l_{\text{flex}}^2}{(\Delta l)^2} p(l_{\text{dsDNA}}, l_{\text{ssDNA}}, l_{\text{linker}}). \qquad [S12]$$

In the first equality, Δl represents the a length scale to transform the flexibility imparted by the Fabs into a number of microstates as in Eq. 3 and Appendix B. 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 Eqs. S10 and S12 into Eq. 2, the relative probability of the bivalent and monovalent states takes the form

$$
\frac{p_2}{p_1} = \frac{e^{-\beta(E_2-E_1)}dV}{(\Delta l)^2} \frac{l_{\text{flex}}^2}{l_{\text{linker}} l_{\text{dsDNA}}} \sqrt{\frac{3}{16\pi^3 n b_{\text{ssDNA}}^2}} \, e^{-\frac{3(l_{\text{linker}}^2+l_{\text{dsDNA}}^2)}{4n b_{\text{ssDNA}}^2}} \sinh\left(\frac{3 l_{\text{linker}} l_{\text{dsDNA}}}{2n b_{\text{ssDNA}}^2}\right)
$$

$$
\equiv \tilde{\alpha}\tilde{n}(l_{\text{dsDNA}}, l_{\text{ssDNA}}) \tag{S13}
$$

where we have defined the prefactor $\tilde{\alpha} = \frac{e^{-\beta(E_2-E_1)}dV}{(\Delta l)^2}$ containing the unknown constants that are independent of the diFab and Env and $\tilde{n}(l_{\text{dsDNA}}, l_{\text{ssDNA}})$ to be the remaining geometry-dependent terms as per Eq. 7. Using the normalization condition $p_0 + p_1 + p_2 = 1$, we recover Eq. 8 for the % neutralization of HIV-1 for a diFab linked together by dsDNA and ssDNA.

As a final aside, note that the small persistence length of the ssDNA ($\xi_{\text{ssDNA}} = 0.1$ nm ≈ 0.1 bases) of the ssDNA implies that there are at least $n = \frac{12 \text{ bases}}{0.1 \text{ bases}} = 120$ segments in each random walk. In this limit, the ideal chain model and worm-like chain model converge to the same result, and hence working with a more complicated polymer physics framework will not alter the result.

D. Model Parameters and Alternate Characterizations of the Data

In this section, we consider the following three models: (*i*) The model used throughout the main text that includes the effect of ssDNA with a persistence length of $\xi_{\text{ssDNA}} = 0.1 \text{ nm} \approx 0.1 \text{ bases}$ of free ssDNA, (*ii*) the model where ssDNA is included in the linker with the persistence length $\xi_{\text{ssDNA}} = 1.5 \text{ nm} = 2.3 \text{ bases}$ of free ssDNA in solution, and (*iii*) the model where ssDNA is completely ignored within the linker. For each model, we list the best-fit parameter values and show how well the model compares to the experimental fits of the data.

All fit parameters were inferred using NonlinearModelFit in Mathematica, where the (logarithms of the) model IC_{50} were fit to the experimentally measured values. Logarithms were used for fitting stability and to prevent the large IC_{50} values from disproportionately influencing the results. Rather than using the entire length titration data in Fig. 1C, fitting was done using the five diFab lengths *d*=50, 56, 60, 64, 68 around the optimal diFab length to ensure that the diFab length data does not dominate the fit results (see Supplementary Mathematica notebook where the fitting and plots are reproduced).

In addition to the parameters shown, all models converted between length and base pairs of dsDNA using $l_{\text{dsDNA}} = d(0.34\frac{\text{nm}}{\text{bp}})$ and for ssDNA using $l_{\text{ssDNA}} = s\left(0.64\frac{\text{nm}}{\text{base}}\right)$. While the value $l_{\text{linker}} = 21 \text{ nm}$ was inferred through fitting, we note that this value could also have been determined without recourse to fitting by noting that the optimal diFab in Fig. 1C had a dsDNA length of *d*=62 corresponds to 21 nm.

Modeling diFab Linkers with a Reduced ssDNA Persistence Length ($\xi_{ssDNA} = 0.08$ nm)

The best fit parameters for this model are given in Table S1. The results of this model are shown in the main text (Fig. 2D, Fig. 3C, Fig. 4C, and Fig. 5C), and we will not reproduce them here.

Description	Parameter	Value	Method Obtained
Dissociation constant of the first diFab arm binding to Env	$K_D^{(1)}$	74 nM	Fab neutralization data (not fit)
Accounts for diFab flexibility from (*i*) Fab $C_H 1$-C_L flexibility with respect to V_H-V_L domains, (*ii*) disorder within the $C_H 1$ C-termini, and (*iii*) the Sulfo-SMCC linker	l_{flex}	1.8 nm	Fit from data
Distance between Fab $C_H 1$ domain C-termini that must be spanned by the dsDNA+ssDNA linker for a bivalently-bound diFab	l_{linker}	21 nm	Fit from data
Persistence length of ssDNA in diFab constructs	ξ_{ssDNA}	0.1 nm	Fit from data
The energetic and entropic increase in relative probability of a bivalently- versus a monovalently-bound diFab (i.e., the avidity effect)	$\tilde{\alpha}\,\tilde{n}(l_{dsDNA}, l_{ssDNA})$	$\tilde{\alpha} = 5.2 \times 10^5$	$\tilde{\alpha}$ fit from data; $\tilde{n}$ was not fit but determined by the diFab composition as per Eqs. 6 and 7

Table S1. Model parameters using ssDNA with the persistence length $\xi_{ssDNA} = 0.08$ nm.

Modeling diFab Linkers with the Persistence Length for Free ssDNA ($\xi_{ssDNA} = 1.5$ nm)
The simplest model of the diFab linker that includes both the dsDNA and ssDNA would characterize the dsDNA as a rigid rod and the ssDNA as an ideal chain with the persistence length of ξ_{ssDNA} =2-6 bases determined for free ssDNA (43, 44).

This model disagrees with the data in several ways: (*i*) The 12 bases ssDNA flanking the dsDNA in Fig. S5A are composed of a few long segments (with Kuhn length $2\xi_{ssDNA} = 4$ bases) that can enable diFabs whose dsDNA is too short or too long to bivalently bind with a minimum entropic cost. Thus, the model predicts a broader IC$_{50}$ profile than suggested by the data. (*ii*) Trading dsDNA for ssDNA in the rigidity constructs in Fig. S5B incurs a minimal cost in potency. For example, even though the measured IC$_{50}$ value of (*d*=62, *s*=12) is about 100x smaller than that of (*d*=50, *s*=18), this model predicts less than a 2x difference between the two constructs since these twelve nucleotides only represent three additional segments as ssDNA (with a concomitant small increase in the entropic penalty of bivalently binding). (*iii*) Constructs with the same dsDNA length but different ssDNA lengths shown in Fig. S5C are predicted to behave nearly identically because the entropy cost of binding bivalently increases slowly with the number of ssDNA segments (and each segment is composed of 4 bases ssDNA). Parameter values for this model are given in Table S2.

Fig. S5 demonstrates the lack of agreement between the data and a model of ssDNA using $\xi_{ssDNA} = 1.5$ nm = 2 bases (the model is worse for larger ssDNA persistence lengths). Together with the DLS data (Fig. 1D), these results suggest that the ssDNA in the diFab linkers is more compact than free ssDNA.

Description	Parameter	Value	Method Obtained
Dissociation constant of the first diFab arm binding to Env	$K_D^{(1)}$	74 nM	Fab neutralization data (not fit)
Accounts for diFab flexibility from (*i*) Fab C$_H$1-C$_L$ flexibility with respect to V$_H$-V$_L$ domains, (*ii*) disorder within the C$_H$1 C-termini, and (*iii*) the Sulfo-SMCC linker	l_{flex}	1.4 nm	Fit from data
Distance between Fab C-termini that must be spanned by the dsDNA+ssDNA linker for a bivalently-bound diFab	l_{linker}	21 nm	Set to this value to slightly improve model results (see Mathematica notebook)
Persistence length of ssDNA in the diFab	ξ_{ssDNA}	1.5 nm	From literature (43, 44)
The energetic and entropic increase in relative probability of a bivalently- versus a monovalently-bound diFab (i.e., the avidity effect)	$\tilde{\alpha}\,\tilde{n}(l_{dsDNA}, l_{ssDNA})$	$\tilde{\alpha} = 5.1 \times 10^5$	$\tilde{\alpha}$ fit from data; $\tilde{n}$ was not fit but determined by the diFab composition as per Eqs. 6 and 7

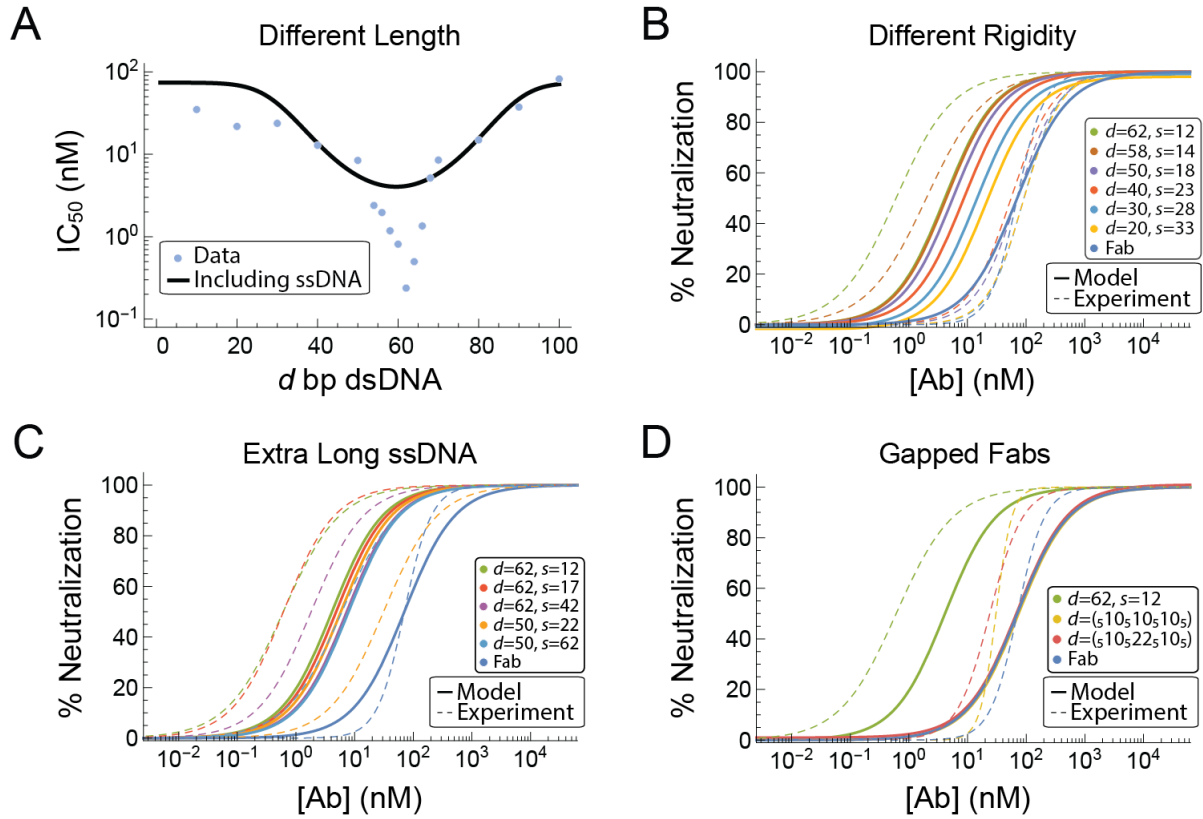**Table S2. Model parameters using ssDNA with the persistence length $\xi_{ssDNA} = 1.5$ nm.**

Figure S5. Modeling the diFab linker with ssDNA included and characterized by its persistence length $\xi_{\text{ssDNA}} = 1.5$ nm in solution. Predicted diFab behavior if the ssDNA in the linker has the same persistence length as free ssDNA in solution. Model results are shown for the analogous plots to (A) Fig. 2D, (B) Fig. 3C, (C) Fig. 4C, and (D) Fig. 5C. Parameter values are given in Table S2.

Modeling diFab Linkers with ssDNA Excluded

Fig. S6 shows the model predictions using Eq. 4, which approximates the ssDNA in the linker as having zero length. Compared to the model we consider in the main text that assumes the ssDNA has a shortened persistence length, this model predicts a sharper length dependence (Fig. S6A) but still captures the general shape of the data. By construction, the extra long constructs (Fig. S6C) in this model depend solely on the length of their dsDNA, and hence the neutralization potencies of diFabs with the same length of dsDNA but different lengths of ssDNA are predicted to be the same.

Note that in Fig. 2D, we showed this model with the slightly sub-optimal $l_{\text{flex}} = 1.8$ nm to match the optimal l_{flex} parameter of the second model that includes ssDNA (Table S1). In Fig. S6, we show the optimal model results when the ssDNA is excluded, with parameters given in Table S3.

Description	Parameter	Value	Method Obtained
Dissociation constant of the first diFab arm binding to Env	$K_D^{(1)}$	74 nM	Fab neutralization data (not fit)
Accounts for diFab flexibility from (*i*) Fab C_H1-C_L flexibility with respect to V_H-V_L domains, (*ii*) disorder within the C_H1 C-termini, and (*iii*) the Sulfo-SMCC linker	l_{flex}	2.6 nm	Fit from data
Distance between Fab C_H1 C-termini that must be spanned by the dsDNA linker for a bivalently-bound diFab	l_{linker}	21 nm	Fit from data
The energetic and entropic increase in relative probability of a bivalently versus a monovalently bound diFab (i.e., the avidity effect)	$\alpha\, n(l_{\text{dsDNA}}, l_{\text{ssDNA}})$	$\alpha = 40$	α fit from data; n was not fit but determined by the diFab composition as per Eq. 3

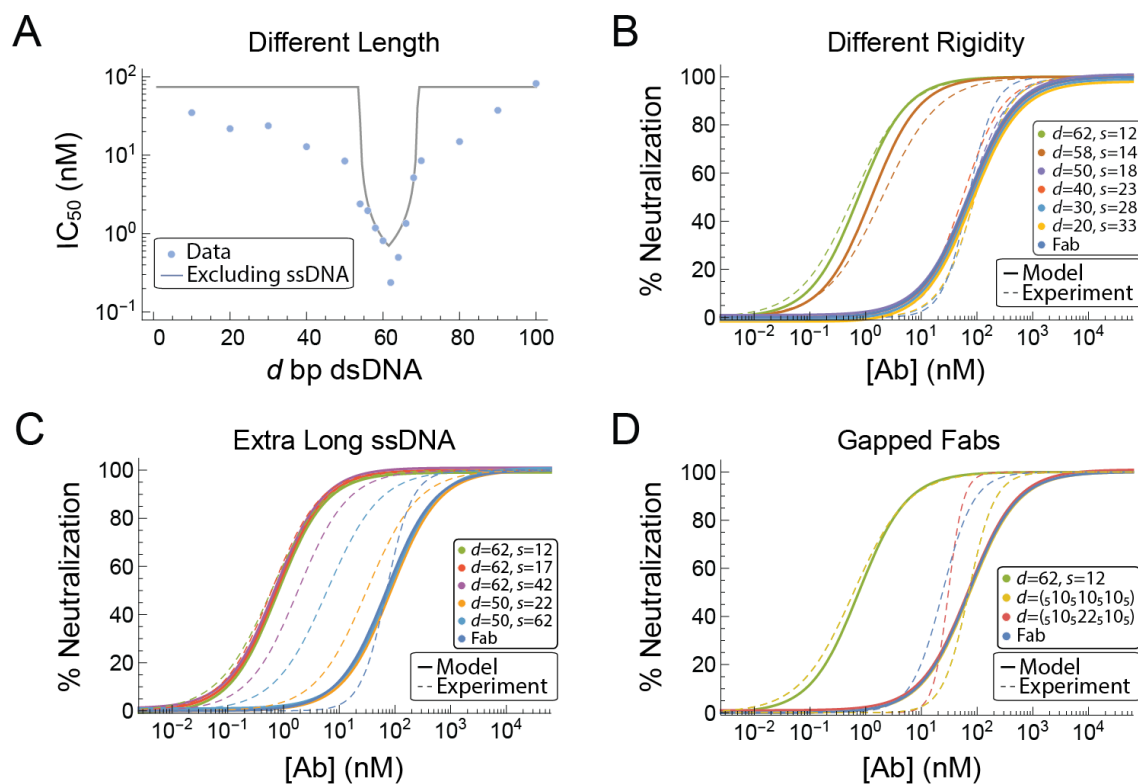**Table S3. Model parameters excluding the ssDNA in the linker.**

Figure S6. Modeling the diFab linker with ssDNA excluded. When diFab behavior is determined solely by its dsDNA content, HIV-1 neutralization is given by Eq. 4. Model results are shown for the analogous plots to (A) Fig. 2D, (B) Fig. 3C, (C) Fig. 4C, and (D) Fig. 5C. Parameter values are given in Table S3.

E. diFab Potencies Against the HIV-1 Strain THRO4156

To verify that the loss in diFab potency when dsDNA was exchanged for ssDNA (Fig. 3) was an inherent property of the linker and not a specific response to the Tier 1B HIV-1 strain 6535.3 used in our neutralization assays, we compared the neutralization potencies of the rigidity diFabs against the Tier 1B HIV-1 6535.3 strain and the Tier 2 HIV-1 strain THRO4156. In both cases, there is a sharp decline in diFab potency when the dsDNA length decreases from d=62 to d≤50.
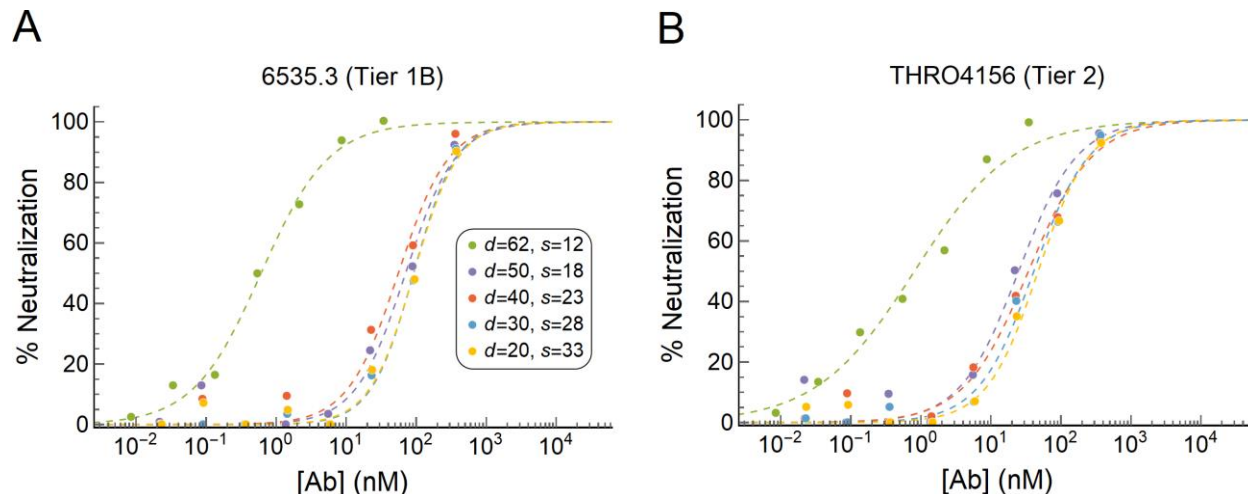

Figure S7. Comparison of Neutralization of Rigidity diFabs against two different HIV-1 Strains. Rigidity diFabs were constructed to keep the length of DNA held constant while ssDNA bases were substituted for dsDNA. (A) Neutralization of the 3BNC60 rigidity diFabs (Fig. 3B) against the Tier 1B HIV-1 strain 6535.3, showing a steep decline in diFab efficacy when the dsDNA length was decreased below d=62 bp. (B) Neutralization of the same diFabs against the Tier 2 HIV-1 strain THRO4156 showing a similar sharp dependence on the dsDNA length.

F. Generalizing the Model to Gapped Fabs and TriFabs

In this section, we discuss how the model of diFab neutralization (Eq. 8) can be generalized to account for either gapped Fabs with ssDNA breaks within the dsDNA segment (Fig. 5A) or triFabs that combine three Fabs to achieve greater avidity effects.

Gapped Fabs represent a Random Walk of dsDNA

Suppose the dsDNA portion of the linker is composed of n_1 segments of d_1 bp dsDNA and one segment of d_2 bp dsDNA flanked by s ssDNA on either side (the constructs in Fig. 5A are ($n_1 = 2$, $d_1 = 10$, $d_2 = 10$, $s = 17$) and ($n_1 = 2$, $d_1 = 10$, $d_2 = 22$, $s = 17$)). The rigidity and extra long constructs have $n_1 = 0$ so that the single dsDNA segment d_2 is flanked by two ssDNA segments of length s that are treated as random walks. For the gapped Fabs, we will consider the two ssDNA segments as well as the n_1 dsDNA segments of length d_1 as random walks about the single dsDNA segment of length d_2. In doing so, we assume that the 5 ssDNA bases interspersed between the dsDNA as free hinges with negligible lengths.

Eq. S9 in Appendix C shows that in the case $n_1 = 0$ where the linker must span the distance $\vec{l}_{\text{linker}}$, the combined ssDNA random walk ($2n$ segments of length b_{ssDNA}) starting at the origin must end on a sphere of radius l_{dsDNA} around $\vec{l}_{\text{linker}}$. 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 $l_{d_2} = d_2 (0.34 \frac{\text{nm}}{\text{bp}})$ around $\vec{l}_{\text{linker}}$, representing the length of the d_2 segment. Using the probability density Eq. S5 for a random walk, the probability that the linker in a gapped Fab will bivalently bind is given by

$$p_{\text{gap}}(l_{\text{dsDNA}}, l_{\text{ssDNA}}, l_{\text{linker}}) = \int_{\vec{x} \in B(\vec{l}_{\text{linker}}, l_{d_2})} \int_{\vec{z} \in \mathbb{R}^3} P_{\text{ssDNA}}(\vec{z}) P_{\text{dsDNA}}(\vec{x} - \vec{z}) d^3\vec{z}\, d^3\vec{x} \qquad [\text{S14}]$$

where

$$P_{\text{ssDNA}}(\vec{z}) = \left(\frac{3}{2\pi b_{\text{ssDNA}}^2 (2n)} \right)^{\frac{3}{2}} e^{-\frac{3z^2}{2b_{\text{ssDNA}}^2 (2n)}} \qquad [\text{S15}]$$

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

$$P_{\text{dsDNA}}(\vec{x} - \vec{z}) = \left(\frac{3}{2\pi l_{d_1}^2 n_1} \right)^{\frac{3}{2}} e^{-\frac{3|\vec{x} - \vec{z}|^2}{2l_{d_1}^2 n_1}} \qquad [\text{S16}]$$

denotes the probability that the endpoint of the dsDNA random walk starting at $\vec{z}$ will end at $\vec{x}$, and $l_{d_1} = d_1 (0.34 \frac{\text{nm}}{\text{bp}})$. Replacing $p \to p_{\text{gap}}$ and $\bar{\Omega}^{2n} \to \bar{\Omega}^{2n+n_1}$ in Eq. S12 yields the desired number of microstates for the bivalent configuration.

Fig. S8 shows how the potency of the optimal (d=62, s=12) construct decreases as the dsDNA segment is broken up into n' pieces of equal length ($n_1 = n' - 1$, $d_1 = d_2 = \frac{62}{n'}$, $s = 12$). Since the 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.
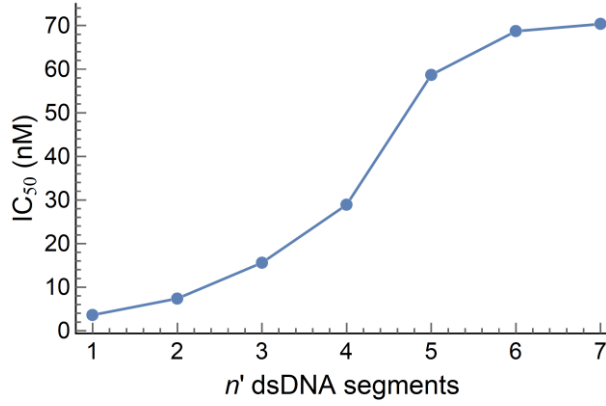
Figure S8. Potency of a gapped diFab made up of 62 bp dsDNA broken into n' equal length pieces. Breaking up the dsDNA increases the entropic cost of bivalent binding, decreasing its potency.

triFabs exhibit Greater Avidity than diFabs

Since the two binding arms in the optimal diFab can decrease its IC_{50} by over 100x compared to the one-armed Fab (21), it is tempting to ask whether a construct with additional Fabs could further reduce the IC_{50}. Here, we outline how our model can be extended to consider the a linear triFab consisting of three Fabs linked in a linear fashion via 62 bp dsDNA segments. The model predicts that optimal triFab will be 100x 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.

As in Eq. 9 (the model that considerers both the dsDNA and ssDNA in the linker), we neglect both the self-intersection of the linker as well as the intersection between the linker and 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 simultaneously binding nearest neighbor arms will be different from binding 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 Fig. 2B for the diFab) are 1, $\frac{[\text{Ab}]}{K_D^{(1)}}$, $\frac{[\text{Ab}]}{K_D^{(1)}} \tilde{\alpha}\, \tilde{n}(l_{\text{dsDNA}}, l_{\text{ssDNA}})$, and $\frac{[\text{Ab}]}{K_D^{(1)}} \tilde{\alpha}^2\, \tilde{n}(l_{\text{dsDNA}}, l_{\text{ssDNA}})^2$ for the states with 0, 1, 2, and 3 Fab arms bound, respectively. Analogous to Eqs. 8 and 9, the probability of binding (and hence neutralizing) a spike is given by

$$p_1 + p_2 + p_3 = \frac{\dfrac{[\text{Ab}]}{K_D^{(1)}} + \dfrac{[\text{Ab}]}{K_D^{(1)}} \tilde{\alpha}\, \tilde{n}(l_{\text{dsDNA}}, l_{\text{ssDNA}}) + \dfrac{[\text{Ab}]}{K_D^{(1)}} \tilde{\alpha}^2\, \tilde{n}(l_{\text{dsDNA}}, l_{\text{ssDNA}})^2}{1 + \dfrac{[\text{Ab}]}{K_D^{(1)}} + \dfrac{[\text{Ab}]}{K_D^{(1)}} \tilde{\alpha}\, \tilde{n}(l_{\text{dsDNA}}, l_{\text{ssDNA}}) + \dfrac{[\text{Ab}]}{K_D^{(1)}} \tilde{\alpha}^2\, \tilde{n}(l_{\text{dsDNA}}, l_{\text{ssDNA}})^2} \quad [\text{S17}]$$

with an IC_{50} given by

$$\text{IC}_{50}^{\text{triFab}} = \frac{K_{\text{D}}^{(1)}}{1 + \tilde{\alpha}\,\tilde{n}(l_{\text{dsDNA}}, l_{\text{ssDNA}}) + \tilde{\alpha}^2\,\tilde{n}(l_{\text{dsDNA}}, l_{\text{ssDNA}})^2}\,. \tag{S18}$$

Substituting in the dsDNA length $l_{\text{dsDNA}} = (62\text{ bp})(0.34\frac{\text{nm}}{\text{bp}})$ from the optimal diFab and assuming that ssDNA strands of lengths $l_{\text{ssDNA}} = (12\text{ bp})(0.64\frac{\text{nm}}{\text{bp}})$ connect the Fabs and the dsDNA, we predict that the triFab should be able to achieve an $\text{IC}_{50}^{\text{triFab}} = 0.008$ nM, representing a 100x improvement over the optimal $\text{IC}_{50}^{\text{diFab}} = 0.8$ nM.

G. Bridge and Linker Sequences for diFabs

Table S4 shows the bridge and linker sequences spanning the two diFab arms in the new constructs created in this work. Sequences were chosen to lack secondary structure and maintain thermal stability (see Methods).

Construct	dsDNA (*d*)	ssDNA (*s*)	DNA Sequence
Rigidity (Fig. 3)	58	14	5-/5AmMC6/TTTTTTTTTTTTCTTTGTTCTTATTCTCTGCTTTTTCTTTCTTTCCTCCTTCTCCCTCTTCTTCCCTTCTTC -3 5-/5AmMC6/AAGAGAGAGAAAAGGAAGAAGGGAAGAAGAGGGAGAAGGAGGAAAGAAAGAAAAAGCAGAGAATAAGAACAA -3
	50	18	5-/5AmMC6/TTTTTTTTTTTTCTTTGTTCTTATTCTCTGCTTTTTCTTTCTTTCCTCCTTCTCCCTCTTCTTCCCTT -3 5-/5AmMC6/AAGAGAGAGAAAAGGAAGAAGGGAAGAAGAGGGAGAAGGAGGAAAGAAAGAAAAAGCAGAGAATAAGA -3
	40	23	5-/5AmMC6/TTTTTTTTTTTTCTTTGTTCTTATTCTCTGCTTTTTCTTTCTTTCCTCCTTCTCCCTCTTCTT -3 5-/5AmMC6/AAGAGAGAGAAAAGGAAGAAGGGAAGAAGAGGGAGAAGGAGGAAAGAAAGAAAAAGCAGAGAA -3
	30	28	5-/5AmMC6/TTTTTTTTTTTTCTTTGTTCTTATTCTCTGCTTTTTCTTTCTTTCCTCCTTCTCCCTC -3 5-/5AmMC6/AAGAGAGAGAAAAGGAAGAAGGGAAGAAGAGGGAGAAGGAGGAAAGAAAGAAAAAGCA -3
	20	33	5-/5AmMC6/TTTTTTTTTTTTCTTTGTTCTTATTCTCTGCTTTTTCTTTCTTTCCTCCTTCT -3 5-/5AmMC6/AAGAGAGAGAAAAGGAAGAAGGGAAGAAGAGGGAGAAGGAGGAAAGAAAGAAA -3
Extra-long (Fig. 4)	62	17	5- /5AmMC6/TTTTTTTTTTTTTTTTTCTTTGTTCTTATTCTCTGCT -3 5- /5Phos/TTTTCTTTCTTTCCTCCTTCTCCCTCTTCTTCCCTTCTTCCT -3 5- /5Phos/GAGAAGGAGGAAAGAAAGAAAAAGCAGAGAATAAGAACAAAG -3 5- /5AmMC6/AAGAGAAGAGAGAGAGAAAAGGAAGAAGGGAAGAAGAGG -3
	62	42	5-/5AmMC6/TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTTTGTTCTTATTCTCTGCT -3 5- /5Phos/TTTTCTTTCTTTCCTCCTTCTCCCTCTTCTTCCCTTCTTCCT -3 5- /5Phos/GAGAAGGAGGAAAGAAAGAAAAAGCAGAGAATAAGAACAAAG-3 5-/5AmMC6/AAGAGAGAGAAAGAGAGAGAAAGAGAGAGAAAGAGAGAGAAAAGGAAGAAGGGAAGAAGAGG -3
	50	22	5- /5AmMC6/TTTTTTTTTTTTTTTTTCTTTGTTCTTATTCTCTGCT -3 5- /5Phos/TTCTTTCTTTCCTCCTTCTCCCTCTTCTTCCCTTC -3 5- /5Phos/GAGAAGGAGGAAAGAAAGAAAGCAGAGAATAAGAA -3 5- /5AmMC6/AAGAGAAGAGAGAGAAAAGGAAGAAGGGAAGAAGAGG -3
	50	62	5-/5AmMC6/TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTTTGTTCTTATTCTCTGCT -3 5- /5Phos/TTCTTTCTTTCCTCCTTCTCCCTCTTCTTCCCTTC -3 5- /5Phos/GAGAAGGAGGAAAGAAAGAAAGCAGAGAATAAGAA -3

			5- /5AmMC6/AAGAGAGAGAAAGAGAGAGAAAGAGAGAGAAAGAGAGAGAA AAGGAAGAAGGGAAGAAGAGG -3
Gapped (Fig. 5)	$_510_510_510_5$	17	5- /5AmMC6/TTTTTTTTTTTTCTTTGTTCTTATTCT -3 5- /5Phos/TTTCTTCTCCCTCTTCTTCCCTTCT -3 5- /5Phos/GGAGAAGAAAAGCAGAGAATAAGAA -3 5- /5AmMC6/AAGAGAGAGAAAAAGGAAGAAGGGAAG -3
	$_510_522_510_5$	17	5- /5AmMC6/TTTTTTTTTTTTCTTTGTTCTTATTCT -3 5- /5Phos/TTTCTTTCTTTCCTCCTTCTCCCTCTTCTTCCCTTCT -3 5- /5Phos/GGAGAAGGAGGAAAGAAAGAAAAGCAGAGAATAAGAA -3 5- /5AmMC6/AAGAGAGAGAAAAAGGAAGAAGGGAAG -3

Table S4. Linker sequences In the diFab constructs.