Supporting Information on: Quantitative Prediction of Multivalent Ligand-Receptor Binding Affinities for Influenza, Cholera and Anthrax Inhibition

Susanne Liese\textsuperscript{1,2} and Roland R. Netz\textsuperscript{1,*}
\textsuperscript{1}Freie Universität Berlin, Department of Physics, 14195 Berlin, Germany
\textsuperscript{2}University of Oslo, Department of Mathematics, 0851 Oslo, Norway
*Email: rnetz@physik.fu-berlin.de

Multivalent Dissociation Constant $K_n$

In the following we derive the multivalent dissociation constant $K_n$ (Eq. 3 in the main text) of an $n$-valent ligand-receptor complex from the partition functions of the bound and unbound ligands.

The multivalent dissociation constant is defined as

$$K_n = \frac{[L][R]}{[LR]}$$  \hspace{1cm} (S1)

with $[L]$ the concentration of unbound multivalent ligands, $[R]$ the concentration of unbound multivalent receptors and $[RL]$ the concentration of ligand-receptor complexes. There is no unambiguous way to distinguish bound and unbound complexes, which is an issue we will discuss further below. If ligand concentration and receptor concentration are dilute, interactions among the ligands as well as among the receptors are negligible and we can apply an ideal gas approximation to relate the dissociation constant to the partition function of the individual constituents

$$K_n = \frac{q_L/V q_R/V}{q_{LR}/V}$$  \hspace{1cm} (S2)

with $q_L$ the partition function of an unbound ligand, $q_R$ the partition function of an unoccupied receptor and $q_{LR}$ the partition function of a ligand-receptor complex. For an $n$-valent system, there are $n$ different binding modes by which ligand and receptors can bind to each other, where the $i$-th binding mode corresponds to $i$ bound ligand units (c.f. Fig.2a in the main text). We denote the partition function of the $i$-th binding mode as $q^{(i)}_{LR}$. The partition function $q_{LR}$ follows as a sum over all $q^{(i)}_{LR}$

$$q_{LR} = \sum_{i=1}^{n} q^{(i)}_{LR}$$  \hspace{1cm} (S3)

According to Eq.S3 we consider a ligand as bound, if at least one ligand unit is bound to a binding pocket. This is a meaningful definition for a large multivalent ligand that competes with a multivalent target for the binding to a multivalent receptor. An influenza virus that targets sialic acid on the cell surface is an example for such a receptor-target system. If the ligand is large enough, it sterically shields the unoccupied binding pockets in addition to the binding pockets that are occupied by ligand units. Hence, in such a system, it is sufficient to bind only one ligand unit to effectively cover the entire receptor.

In the following, we first discuss the partition function of an unbound ligand $q_L$ and subsequently the partition function $q^{(i)}_{LR}$.

The partition function of an unbound ligand $q_L$ reads

$$q_L = V \cdot 8\pi^2 \cdot \Omega_{LU}^n \cdot \prod_{j=1}^{n} \int dr_j P_{str}^{bulk}(r_j),$$  \hspace{1cm} (S4)

with $V$ the system volume, $8\pi^2$ the angular space available to a rigid body that can rotate around all three axes in space, $\Omega_{LU}$ the angular space available to each ligand unit when it is in bulk and thus far away from the receptor surface and $P_{str}^{bulk}(r_j)$ the probability that the $j$-th linker extends along the end-to-end vector $r_j$ (schematically shown in Fig.S1a). We impose that $P_{str}^{bulk}$ is normalized, such that $q_L$ simplifies to

$$q_L = V \cdot 8\pi^2 \cdot \Omega_{LU}^n$$  \hspace{1cm} (S5)

To derive an expression for the partition function $q^{(i)}_{LR}$ of a ligand-receptor complex in the $i$-th binding mode, we denote as $\{i\}$ the set of indices of $i$ bound ligand units and as $\{\{i\},n\}$ the set of combinations to bind $i$ ligand units to an $n$-valent receptor. The
partition function $q_{LR}^{(i)}$ factorizes into the partition function of the receptor $q_R$, which includes all internal degrees of freedom of the receptor that remain unchanged during the binding process, and the integrals over the ligand position and orientation

$$q_{LR}^{(i)} = q_R \cdot \int dr_0 \int d\omega_0 \sum_{\{i,n\}} \left( \prod_{j \in \{i\}} \int \Omega_{bp} V_{bp} \int V_{bp} \int d\omega_{r_j} \int P^\text{bound}_{str}(r_j) m^i \cdot \prod_{j \notin \{i\}} \Omega_{LU} \int dr_j P^\text{im}_{str}(r_j) \right) . \quad (S6)$$

with $r_0$ the position of the core midpoint relative to the receptor midpoint and $\omega_0 = \omega_0(\phi, \psi, \theta)$ the Euler angles by which the core can rotate (schematically shown in Fig.S1b). The number of equivalent binding pockets per receptor subunit is denoted as $m$. As we discuss in the main text, $m=1$ for the trivalent hemagglutinin receptor and pentavalent cholera toxin, but $m=2$ for the heptavalent anthrax receptor. The angular space of a bound ligand unit is denoted as $\Omega_{bp}$. The position of each bound ligand unit is confined to the volume of the binding pocket $V_{bp}$. $P^\text{bound}_{str}(r_j)$ describes the stretching probability of a bound linker, where $r_j$ is a vector that connects the $j$-th ligand unit and the $j$-th ligand core corner, and $P^\text{im}_{str}(r_j)$ is the stretching probability of an unbound linker. Note that unlike $P^\text{bulk}_{str}$, which describes the stretching probability of a linker that connects to a ligand that is not bound to a receptor, $P^\text{bound}_{str}(r_j)$ and $P^\text{im}_{str}(r_j)$ take into account that the linker cannot penetrate the receptor surface.

We here assume that the linker is much longer than the extension of the binding pocket ($r_j \gg V_{bp}^{1/3}$) and hence, the stretching probability is approximately independent of the position inside the binding pocket. Furthermore, each bound ligand unit gains the binding free energy $\Delta G$. The factor $\beta$ denotes the inverse thermal energy. The angular space available to each unbound ligand unit is $\Omega_{LU}$, equivalent to the case of an unbound ligand (Eq.S5). Analytic expressions for $P^\text{bound}_{str}$ and $P^\text{im}_{str}$ are derived further below.

Based on Eq.S2, Eq.S3, Eq.S5 and Eq.S6 the dissociation constant of an $n$-valent ligand is written as

$$K_n = \left[ \sum_{i=1}^n \int dr_0 \int d\omega_0 \frac{1}{8\pi^2} \sum_{\{i,n\}} \left( \prod_{j \in \{i\}} \Omega_{bp} V_{bp} e^{-\beta \Delta G} P^\text{bound}_{str}(r_j) m^i \cdot \prod_{j \notin \{i\}} \Omega_{LU} \int dr_j P^\text{im}_{str}(r_j) \right) \right]^{-1}. \quad (S7)$$

For monovalent ligands and monovalent receptors ($n=1$, $m=1$) the integration over the core position $r_0$ and orientation $\omega_0$ as well as the integration over the linker end-to-end vector $r_j$ and the linker stretching probabilities can be performed. Furthermore, the angular space of the ligand unit $\Omega_{LU}$ is given by the angular space of a freely rotating rigid body, $\Omega_{LU} = 8\pi^2$. Hence, Eq.S7 simplifies to

$$K_1 = \frac{8\pi^2 e^{\beta \Delta G}}{\Omega_{bp} V_{bp}} . \quad (S8)$$

Inserting Eq.S8 into Eq.S7 we obtain the dissociation constant $K_n$ as

$$K_n = \left[ \sum_{i=1}^n \frac{C_n^{(i)} m^i}{K_1^{\Omega_{LU}} \Omega_{LU}} \right]^{-1} = \left[ \sum_{i=1}^n \frac{C_n^{(i)} m^i}{K_1^{\Omega_{LU}}} \right]^{-1} . \quad (S9)$$

with $C_n^{(i)} = \int dr_0 \int d\omega_0 \sum_{\{i,n\}} \left( \prod_{j \in \{i\}} P^\text{bound}_{str}(r_j) \cdot \prod_{k \notin \{i\}} \int dr_k P^\text{im}_{str}(r_k) \right) . \quad (S10)$

For a fully bound ligand, i.e. if we neglect all terms for $i < n$, the dissociation constant $K_n$ is given by Eq. 3 from the main text

$$K_n = \frac{K_n^{\Omega_{LU}}}{C_n m^n} = \frac{K_n^{\Omega_{LU}}}{C_n m^n} . \quad (S11)$$

with $C_n := C_n^{(n)} = \int dr_0 \int d\omega_0 \frac{1}{8\pi^2} \sum_{\{n,n\}} P^\text{bound}_{str}(r_j)$.

This approximation will be shown to be accurate further below, where we discuss the impact of partially and fully bound ligands. If the angular space $\Omega_{LU}$ is smaller than the angular space within the binding pocket, $\Omega_{LU} < \Omega_{bp}$, then the angular space of a bound ligand unit is equal to $\Omega_{LU}$. In other words, $\Omega_{bp}$ in Eq.S6 should be replaced by $\Omega_{LU}$ and consequently, $\Omega_{LU}$ in Eq.S11 should be replaced by $\Omega_{bp}$. So, in writing Eq.S11 we assume that the angular space in the binding pocket $\Omega_{bp}$ is smaller than $\Omega_{LU}$.

**Cooperativity Factor $C_n$ for a Fully Bound Ligand**

In the following, we derive an analytic expression for the cooperativity factor $C_n$. First, we discuss the steric repulsion between the linker and the receptor that arises because the linker cannot penetrate the receptor surface. We quantify this effect by deriving
the stretching probability for a flexible polymer close to an impenetrable planar wall. On large scales and neglecting self avoidance effects, a flexible polymer can be described as a Gaussian chain and is characterized by the average end-to-end distance \( r_{\text{ete}} \). The free energy \( F(r) \) in dependence of the end-to-end distance \( r \) reads\(^{1,2}\)

\[
F(r) = \frac{3}{2} \left( \frac{r}{r_{\text{ete}}} \right)^2 k_B T, \tag{S13}
\]

with \( k_B T \) the thermal energy. By construction, the free energy in Eq.S13 leads to an average squared end-to-end distance \( \langle r^2 \rangle = r_{\text{ete}}^2 \) in three dimensions. From the free energy, we obtain the stretching probability \( P_{\text{str}}^{\text{bulk}} \), in the absence of the receptor surface, as

\[
P_{\text{str}}^{\text{bulk}}(z_j, \rho_j) = \frac{\exp \left[ -\beta F(r_j) \right]}{4\pi \int_0^\infty \! dr_j r_j^2 \exp \left[ -\beta F(r_j) \right]} = \left( \frac{3}{2\pi r_{\text{ete}}^2} \right)^{3/2} \exp \left[ \frac{3}{2} \left( \frac{\rho_j^2 + z_j^2}{r_{\text{ete}}^2} \right) \right], \tag{S14}
\]

with \( \rho_j \) and \( z_j \) the radial and the \( z \) components of the end-to-end vector \( r_j \).

**Figure S1:** a) Schematic picture of the linker end-to-end vectors for an unbound ligand. b) Schematic picture of the relative position and orientation of the core with respect to the receptor. The vector from the receptor midpoint to the first binding pocket, \( d_{\text{rec,1}} \), as well as the vector from the core midpoint to the core edge, \( d_{\text{core,1}} \), are shown. The vector that connects the receptor and core midpoints is denoted as \( r_0 \). The rotation of the core around the \( x \), \( y \), \( z \)-axis is described by the angles, \( \phi \), \( \theta \), \( \psi \). c) Schematic picture of the end-to-end vectors of the linkers for a fully bound ligand. d) Image construction for a flexible polymer at an impenetrable wall. For every path that starts at \( z_s \), ends at \( z_e \) and penetrates the wall, there is an equivalent path that starts at \( -z_s \) and ends at \( z_s \). e) Schematic picture of the position and height \( z_s \) of the first linker segment, which is approximated by a freely rotating rigid rod of length \( a_0 \). f) Ligands with long linkers and small cores (top) exhibit a large angular space \( \Omega_a \approx 8\pi^2 \) and all ligand units can reach all binding pockets. The rotation of ligands with short linkers and large cores (bottom) are severely restricted (\( \Omega_a \ll 8\pi^2 \)) and ligand units can reach only the closest binding pocket.

**Steric Repulsion:** At an impenetrable planar wall (Fig.S1d), the stretching probability \( P_{\text{str}}^{\text{im}} \) of a linker that starts at height \( z_s \) and ends at height \( z_e \) follows from an image construction as\(^3\)

\[
P_{\text{str}}^{\text{im}}(z_s, z_e, \rho) = P_{\text{str}}^{\text{bulk}}(z_e - z_s, \rho) - P_{\text{str}}^{\text{bulk}}(z_e + z_s, \rho), \tag{S15}
\]

with \( \rho \) the linker vector component parallel to the surface of the wall. The second term in Eq.S15 ensures that \( P_{\text{str}}^{\text{im}} \) vanishes in the limit \( z_s = 0 \). To regularize Eq.S15 as \( z_s \to 0 \) we treat the first linker segment as a rigid rod of length \( a_0 \) (c.f Fig.S1e). We set \( a_0=0.4 \text{nm} \), which is equivalent to the monomer length of PEG as determined from MD simulations.\(^4\) The probability that the first rigid segment reaches a height \( z_s \) reads

\[
P_{\text{rod}}(z_s) = \frac{4}{\pi} \sqrt{a_0^2 - z_s^2}/a_0^3, \tag{S16}
\]

which is normalized as \( \int_0^{a_0} dz_s P_{\text{rod}}(z_s) = 1 \). The stretching probability \( P_{\text{str}}^{\text{bound}} \) of a bound linker is obtained by integrating Eq.S15 over \( z_s \) weighted with \( P_{\text{rod}}(z_s) \). In the limit that the \( z \)-position \( z_s \) is much smaller the the average end-to-end distance, \( z_s/r_{\text{ete}} \ll 1 \),
\[ p_{\text{bound}}^\text{str}(z_j, \rho_j) = \int_0^{z_0} dz_0 p_{\text{rod}}(z_0) p_{\text{str}}^\text{im}(z_j, z_0, \rho_j) \]
\[ \approx \left( \frac{3}{2\pi r_{\text{ete}}} \right)^{3/2} \exp \left[ -\frac{3\rho_j^2 + z_j^2}{2r_{\text{ete}}} \right] \frac{8a_0 z_j}{\pi r_{\text{ete}}} \]
\[ \approx \left( \frac{3}{2\pi r_{\text{ete}}} \right)^{3/2} \exp \left[ -\frac{3\rho_j^2 + z_j^2}{2r_{\text{ete}}} \right] \frac{8a_0 z_0}{\pi r_{\text{ete}}} \]

For a fully bound ligand the heights of the ligand core corners \( z_j \) are similar to the height of the core midpoint \( z_0 \). We therefore approximate \( z_j \) in the exponential prefactor in Eq.S17 by \( z_0 \).

**Angular Space and Conformational Weight:** Inserting Eq.S17 into Eq.S12, we see that the cooperativity factor \( C_n \) depends on the sum over the squared end-to-end distances \( \sum_j (\rho_j^2 + z_j^2) = \sum_j r_j^2 \). It is therefore useful to derive an expression for the squared end-to-end distances summed over all linkers. The position of the core corners is obtained by rotating \( \mathbf{r}_{\text{core},j} \), the vector from the midpoint of the core to the \( i \)-th corner, by \( \phi, \theta, \psi \) around the \( x, y, z \) axis. Subsequently, the core midpoint is moved by \( r_0 \) relative to the midpoint of the receptor, as schematically depicted in Fig.S1b. The vector from the \( j \)-th corner of the ligand core to the \( k \)-the binding pocket reads

\[ \mathbf{r}_j(k) = \left( \begin{array}{c} d_{\text{core}} \cos \left( \frac{2\pi k(j)}{n} \right) \\ d_{\text{core}} \sin \left( \frac{2\pi k(j)}{n} \right) \end{array} \right) \left[ \begin{array}{c} \rho_0 \cos (\phi_0) \\ \rho_0 \sin (\phi_0) \end{array} \right] + R_z(\phi) \cdot R_y(\theta) \cdot R_x(\psi) \left( \begin{array}{c} d_{\text{core}} \cos \left( \frac{2\pi j}{n} \right) \\ d_{\text{core}} \sin \left( \frac{2\pi j}{n} \right) \end{array} \right), \]

with \( R_{x,y,z} \) the rotation matrix around the \( x, y, z \) axis and \( \rho_0, \phi_0, z_0 \) the components of \( \mathbf{r}_0 \) in cylindrical coordinates. For a divergent ligand the rotation around \( x \) is omitted, due to the rotational symmetry of the core.

The sum \( \sum_{j=1}^n \mathbf{r}_j^2(k) \) cannot be evaluated in closed form for arbitrary combinations of binding pockets and ligand units, i.e. for arbitrary \( k(j) \). However, for \( k = j \), the sum over \( \mathbf{r}_j^2(k) \) can be performed exactly and yields

\[ \sum_{j=1}^n \mathbf{r}_j^2(k) = \begin{cases} \frac{2}{n} \left( d_{\text{dec}} - d_{\text{core}} \right)^2 + \rho_0^2 + z_0^2 - d_{\text{dec}} d_{\text{core}} (\cos \phi + \cos \theta \cos \phi - 2) & n = 2, \\
\frac{n}{n} \left( d_{\text{dec}} - d_{\text{core}} \right)^2 + \rho_0^2 + z_0^2 - d_{\text{dec}} d_{\text{core}} (\sin \phi \sin \theta \sin \psi + \cos \phi \cos \psi + \cos \theta \cos \phi - 2) & n > 2. \end{cases} \]

To discuss the impact of the restriction to ligand unit-binding pocket combinations with \( k = j \), we consider two limiting cases.

1. A small, point-like core \( (d_{\text{core}} \ll d_{\text{dec}}) \): If the core is very small, all permutations of ligand unit - binding pocket combinations are equivalent, since the core has the same distance from all binding pockets (Fig.S2a). The sum over the number of permutations in Eq.S12 in this case yields a factor of \( n! \) and we obtain

\[ C_n = \int d\mathbf{r}_0 \int d\omega_0 \frac{1}{8\pi^2} n! \prod_j p_{\text{str}}^\text{bound} \left( \mathbf{r}_j(k = j) \right) \quad \text{if} \quad d_{\text{core}} \ll d_{\text{dec}}. \]

2. A ligand core that is similar in size to the receptor \( (d_{\text{core}} \approx d_{\text{dec}}) \): If the ligand core is similar in size to the receptor, the condition \( k = j \) in Eq.S19 means that consecutive ligand units bind to consecutive binding pockets (Fig.S2b). This condition neglects permutations, for which the linkers are considerably stretched (schematically shown in Fig.S2c). Neglecting these deformations is justified, since extended linkers result in a small stretching probability \( p_{\text{bound}}^\text{str} \) and therefore do not contribute significantly to the cooperativity factor \( C_n \). Further below, we show that this approximation is accurate by comparing the approximate expression for \( C_n \) (Eq.S22) with a numerical Monte-Carlo integration of Eq.S10 and S12. In the limit \( d_{\text{core}} \approx d_{\text{dec}} \), there are \( n \) permutations of ligand unit - binding pocket combinations that are equivalent to the condition \( k = j \) (Fig.S2b), hence the cooperativity factor reads

\[ C_n = \int d\mathbf{r}_0 \int d\omega_0 \frac{1}{8\pi^2} n \prod_j p_{\text{str}}^\text{bound} \left( \mathbf{r}_j(k = j) \right) \quad \text{if} \quad d_{\text{core}} \approx d_{\text{dec}}. \]

To combine these two limits into a single expression, we introduce a permutation factor \( \Pi(n) \), which describes the number of
permutations by which the $n$ ligand units can bind to the $n$ binding pockets. The cooperativity factor $C_n$ then factorizes as

$$C_n = \Pi(n) \frac{\Omega_c}{8\pi^2} Q_c,$$

with $\Omega_c = \left\{ \begin{array}{ll} \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin(\theta) \int_0^{2\pi} d\psi \exp \left[ 3 \frac{d_{\text{core}} d_{\text{rec}}}{r_{\text{ete}}} (\cos \phi + \cos \theta \cos \phi - 2) \right] & \text{for } n = 2, \\ \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin(\theta) \int_0^{2\pi} d\psi \exp \left[ \frac{3n}{2\pi} \frac{d_{\text{core}} d_{\text{rec}}}{r_{\text{ete}}} (\sin \phi \sin \theta \sin \psi + \cos \phi \cos \psi + \cos \theta \cos \psi - 2) \right] & \text{for } n > 2, \end{array} \right.$$

and $Q_c = e^{-\frac{3n}{2} \left( \frac{d_{\text{core}} - d_{\text{rece}}}{r_{\text{ete}}} \right)^2} \int_V d\rho_0 \left( P_{\text{bound}}(r_0) \right)^n$.

$$= e^{-\frac{3n}{2} \left( \frac{d_{\text{core}} - d_{\text{rece}}}{r_{\text{ete}}} \right)^2} \frac{4\pi}{3} \int_0^\infty d\rho_0 \int_0^{\infty} d\zeta_0 \left( \frac{3}{2\pi r_{\text{ete}}} \right)^{3n/2} \exp \left[ -\frac{3n}{2} \sum_{j=1}^{n} r_{j}^2 (\rho_0, \zeta_0) \right] \left( \frac{8\pi \zeta_0}{\pi r_{\text{ete}}} \right)^n.$$

$$= e^{-\frac{3n}{2} \left( \frac{d_{\text{core}} - d_{\text{rece}}}{r_{\text{ete}}} \right)^2} \sum_{j=1}^{n} r_{j}^2 (\rho_0, \zeta_0) \left( \frac{8\pi \zeta_0}{\pi r_{\text{ete}}} \right)^n.$$

The terms $\Omega_c$ and $Q_c$ are the angular and conformational weights, which account for the restriction of the core rotational degrees of freedom as well as the restriction of the core positional degrees of freedom due to the linker stretching.

As we discuss in detail further below, the angular weight $\Omega_c$ exhibits two limits, with $\Omega_c \rightarrow 8\pi^2$ for small, point-like cores and $\Omega_c \rightarrow 0$ for large cores. To obtain an expression that interpolates between $\Pi(n) = n$ for large cores and $\Pi(n) = n!$ for point-like cores, we approximate the number of permutations as

$$\Pi(n) = n \frac{8\pi^2 - \Omega_c}{8\pi^2} + n \frac{\Omega_c}{8\pi^2}.$$

For all figures shown in the main text, Eq.S26, Eq.S25 and Eq.S23 are inserted into Eq.S22, and the angular weight $\Omega_c$ is evaluated numerically.

**Analytic Approximation:** In the case of a ligand core that is similar in size to the receptor, $d_{\text{core}}/d_{\text{rec}} \approx 1$, as well as for very small, point-like cores, $d_{\text{core}}/d_{\text{rec}} \ll 1$, an analytic approximation for the cooperativity factor $C_n$ can be derived, as we show now. If the ligand core is much smaller than the linker, $d_{\text{core}}/r_{\text{ete}} \ll 1$, the restriction of the rotational degrees of freedom of the
core is negligible, i.e. in the limit $3n_{rec}d_{core}/(2r_{ete}^2) \to 0$, the integration over the Euler angles in Eq.S23 simplifies and we obtain $\Omega_c = 8\pi^2$. If the linker size is much shorter than the receptor size and the core size, or equivalently $3n_{rec}d_{core}/(2r_{ete}^2) \gg 1$, the integration over the Euler angles in Eq.S23 is performed by a saddle-point approximation around $\theta = \psi = 0$. The two limits of $\Omega_c$ are summarized as

$$
\Omega_c \approx \begin{cases} 
8\pi^2, & n = 2 \text{ and } \frac{r_{ete}^2}{d_{core}^2} \gg 6(2\pi)^{1/3} \\
2\left(\frac{3}{2\pi}\right)^{3/2}\left(\frac{d_{core}}{d_{rec}}\right)^3\left(\frac{d_{core}}{d_{rec}} - 1\right)^3, & n = 2 \text{ and } \frac{r_{ete}^2}{d_{core}^2} \leq 6(2\pi)^{1/3} \\
\frac{2^{3/2}}{\pi n} \left(\frac{2}{\pi}\right)^2\left(\frac{d_{core}^4}{d_{core}^2 - d_{rec}^2}\right)^4\left(\frac{d_{core}^4}{d_{core}^2 - d_{rec}^2} - 1\right)^4, & n > 2 \text{ and } \frac{r_{ete}^2}{d_{core}^2} > \frac{3n\sqrt{\pi}}{2^{3/4}} \\
8\pi^2, & \text{if } r_{ete} \gg \sqrt{\frac{d_{core}^2}{\pi n}}\end{cases} 
$$

(S27)

According to the discussion leading to Eq.S26, the two limits of the number of permutations $\Pi(n)$ are

$$
\Pi(n) \approx \begin{cases} 
n!, & \frac{d_{core}}{d_{rec}} \ll 1 \\
n, & \frac{d_{core}}{d_{rec}} \approx 1.
\end{cases} 
$$

(S28)

Inserting Eq.S25, S27 and S28 into Eq.S22, we obtain the following expression for the cooperativity factor $C_n$

$$
C_n \approx \begin{cases} 
2^{-\frac{2}{3n}} \left(\frac{3}{2\pi}\right)^{3/2} \left(\frac{2^6}{3\pi^2}\right)^2 \left(\frac{d_{core}}{d_{rec}}\right)^2 \left(\frac{d_{core}}{d_{rec} - d_{ete}}\right)^{2}, & n = 2 \text{ and } \frac{d_{core}}{d_{rec}} \ll 1 \\
2\left(\frac{3}{2\pi}\right)^{3/2} \left(\frac{d_{core}}{d_{rec}}\right)^{-3/2} \left(\frac{d_{core}}{d_{rec} - d_{ete}}\right)^{-2} \left(\frac{d_{core} - d_{core}}{d_{core} - d_{ete}}\right)^{2} e^{-\frac{1}{2} \left(\frac{d_{core} - d_{core}}{d_{core} - d_{ete}}\right)^2}, & n = 2 \text{ and } \frac{d_{core}}{d_{rec}} \approx 1 \\
n^{1/2} n^{-\frac{1}{2}} \left(\frac{2}{\pi}\right)^{1/2} \left(\frac{d_{core}}{d_{rec} - d_{ete}}\right)^{2}, & n > 2 \text{ and } \frac{d_{core}}{d_{rec}} \ll 1 \\
n^{3/2} \left(\frac{2}{\pi}\right)^{\frac{3}{2}} \left(\frac{d_{core}}{d_{rec} - d_{ete}}\right)^{2} \left(\frac{d_{core} - d_{core}}{d_{core} - d_{ete}}\right)^{2} \left(\frac{d_{core} - d_{core}}{d_{core} - d_{ete}}\right)^{2}, & n > 2 \text{ and } \frac{d_{core}}{d_{rec}} \approx 1.
\end{cases} 
$$

(S29)

Maximizing Eq.S29 with respect to $r_{ete}$, we find the following optimal linker length $r_{ete}^{opt}$

$$
r_{ete}^{opt} \approx \begin{cases} 
\sqrt{\frac{2}{3} d_{core}}, & n = 2 \text{ and } \frac{d_{core}}{d_{rec}} \ll 1 \\
\sqrt{\frac{2}{3}} |d_{core} - d_{rec}|, & n = 2 \text{ and } \frac{d_{core}}{d_{rec}} \approx 1 \\
\frac{3n - 3}{3n - 3} d_{core}, & n > 2 \text{ and } \frac{d_{core}}{d_{rec}} \ll 1 \\
\frac{3n - 3}{3n - 3} |d_{core} - d_{rec}|, & n > 2 \text{ and } \frac{d_{core}}{d_{rec}} \approx 1.
\end{cases} 
$$

(S30)

Inserting Eq.S30 and Eq.S29 into Eq. 6 in the main text, the monovalent dissociation constant as a function of the enhancement
In Fig.S3 the monovalent dissociation $K_1$ for $\alpha=1$ and the optimal linker length $r_{\text{opt}}$ based on the numerical evaluation of Eq.S22 is compared with the analytic approximations Eq.S30 and S31. In the respective limits the analytic approximations agree well with the numerical results.

![Figure S3](image-url)

**Figure S3:** Enhancement diagram and optimal linker length $r_{\text{opt}}$ based on the numerical evaluation of Eq.S22 (solid lines) and given by the analytic approximations Eq.S31d (dotted lines) and Eq.S31e (dashed lines). The following input parameters are used: a) $n=3$, $d_{\text{rec}}=2.6\text{nm}$, $\omega_{LU} = 0.03$, $m=1$ b) $n=5$, $d_{\text{rec}}=2.5\text{nm}$, $\omega_{LU} = 0.03$, $m=1$ c) $n=7$, $d_{\text{rec}}=3.5\text{nm}$, $\omega_{LU} = 0.03$, $m=2$.

**Validation of the permutation factor $\Pi(n)$:** The only difference between Eq.S12 and Eq.S22 arises from the introduction of the permutation factor $\Pi(n)$, which describes the number of equivalent permutations by which $n$ ligand units can bind to $n$ binding pockets. To validate the approximation for $\Pi(n)$ in Eq.S26, we show in Fig.S4 the dissociation constants $K_5$ and $K_7$ defined in Eq.S11 for cholera toxin and the anthrax receptor based on Eq.S22 (solid lines) as well as based on a numerical evaluation of Eq.S12 (dashed lines). The numerical integration Eq.S12 over the core position $r_0$ and the core angles $\omega_{LU}$ is done using Monte Carlo integration techniques. The experimental data as well as the model parameters, $K_1$, $d_{\text{core}}$, $d_{\text{rec}}$, $\omega_{LU}$ and $m$ are the same as...
in Fig.4 in the main text. The approximate treatment according to Eq.S22 agrees satisfactorily with the full numerical integration of the multidimensional integral Eq.S12. For the pentavalent cholera toxin, the approximation Eq.S22 agrees slightly better with the experimental results than the numerical evaluation of Eq.S12, presumably because the approximation of the permutation factor \( \Pi(n) \) in Eq.S26 effectively corrects for volume exclusion effects between neighboring linkers, which is not included in Eq.S12.

![Figure S4](image)

**Figure S4:** Comparison of the experimental dissociation constant with the theoretical model based on the approximation for \( C_n \) according to Eq.S22 (solid lines) and based on the numerical evaluation of \( C_n \) in Eq.S12 (dashed lines). a) Heptavalent ligand-receptor pair. b) Pentavalent ligand receptor pair. The larger core size (\( d_{\text{core}}=0.8\text{nm} \)) is shown in blue, while the smaller core size (\( d_{\text{core}}=0.3\text{nm} \)) is shown in red.

**Enhancement Diagram**

Eq. 6 in the main text defines the monovalent dissociation constant that is needed to achieve an enhancement in binding efficiency by a factor of \( \alpha \). In Fig.S5a we show the enhancement diagram for tri-, penta- and heptavalent ligand-receptor pairs for four different values of \( \alpha \).

![Figure S5](image)

**Figure S5:** a) Enhancement diagram for tri-, penta- and heptavalent ligand-receptor pairs for four different enhancement factors \( \alpha \). The monovalent dissociation constants for hemagglutinin (\( K_1=2.5\text{mM} \)), cholera toxin (\( K_1=5\text{mM} \)) and anthrax (\( K_1=4\text{mM} \)) are shown as dashed, black lines. b) Enhancement diagram for tri-, penta- and heptavalent ligand-receptor pairs for enhancement factor \( \alpha=1 \) for optimal and non-optimal linker lengths. c) Optimal rescaled linker length \( r_{\text{ete}}^{\text{opt}}/[d_{\text{core}} - d_{\text{rec}}] \). The rescaled linker length that would equal the geometric separation between ligand core corners and receptor units is indicated by a dashed horizontal line.
In accordance with the receptor structures shown in Fig. 1a in the main text, we use the following values for \( m \), the number of binding pockets per receptor subunit, \( n=1 \) (\( n=3 \), \( n=5 \)), and \( n=2 \) (\( n=7 \)). From a fit of the experimental data to our theoretical model, shown in Fig. 4 a and b in the main text, we determined the angular restriction factor \( \omega_{LU} = 0.03 \) (\( n=5 \) and \( n=7 \)). For trivalent ligands we use the same angular restriction factor, \( \omega_{LU} = 0.03 \). The linker length is optimized, such that \( K_n \) becomes minimal. The monovalent dissociation constants for hemagglutinin \((K_1 = 2.5 \text{ mM})\), cholera toxin \((K_1 = 5 \text{ mM})\) and anthrax \((K_1 = 4 \text{ mM})\) are shown as dashed horizontal lines. In all three cases a multivalent ligand that has the same binding affinity as the monovalent counterpart, corresponding to an enhancement factor of unity \( \alpha = 1 \), is obtained for a wide range of core sizes. However, to achieve an enhancement by several orders of magnitude, the core size has to match the receptor size quite precisely; in particular for the trivalent case \( n = 3 \). If the core size and receptor size are similar, the optimal linker length is slightly larger than the difference between \( d_{\text{core}} \) and \( d_{\text{dec}} \) as shown in Fig.S5c.

To study the robustness of the binding enhancement with respect to a non-optimal linker length, Fig.S5b shows \( K_1 \) for \( \alpha = 1 \) for the optimal linker length (solid lines), as well as for linkers that are by 0.3nm too short (dotted lines) or too long (dashed lines). Since the optimal receptor length depends on \( |d_{\text{core}} - d_{\text{rec}}| \), we have to choose a specific value for \( d_{\text{rec}} \) or for \( d_{\text{core}} \) to determine the enhancement diagram for suboptimal linker lengths. Therefore, we set the receptor size to \( d_{\text{rec}} = 2.6 \) (\( n=3 \)), \( d_{\text{rec}} = 2.5 \) (\( n=5 \)) and \( d_{\text{rec}} = 3.5 \text{ nm} \) (\( n=7 \)), corresponding to hemagglutinin, cholera toxin and anthrax. Similar to the results presented in Fig. 5 in the main text, we find that for ligand cores that are similar to the receptor size, the binding enhancement significantly decreases for non-optimal linker length. In fact, the decrease in binding enhancement is much more pronounced for linkers that are shorter than the optimal length, compared to linkers that are longer.

**Impact of Partially Bound Ligands**

To study the impact of partially bound ligands, we show the inverse of the dissociation constant of hepta- and pentavalent ligands in Fig.S6, since the inverse dissociation constant consists of a sum over the individual binding modes. The dissociation constant in Eq.9 is determined by numerically evaluating the cooperativity factors \( C_{LU}^{(i)} \). In analogy to the results shown in Fig.S4, we determine \( C_{LU}^{(i)} \) by numerically integrating Eq.S10 over the core position \( r_0 \), the core angles \( \omega \) and the linker positions \( r_k \) using Monte Carlo integration. The model parameters, \( K_1, d_{\text{core}}, d_{\text{dec}}, \omega_{LU} \) and \( m \) are the same as in Fig.4 in the main text and are identical in all binding modes. Evidently, \( C_{LU}^{(i)} \) changes in each binding mode according to Eq.S10. Fig.S6 shows that the minimum of the dissociation constant, i.e. the maximum of the inverse dissociation constant, is dominated by fully bound ligands. If the linker length is larger or smaller than the optimal value, the contribution of partially bound ligands surpasses the contribution of fully bound ligands. In the three examples studied here, the deviation between the dissociation constant taking all binding modes into account (dotted line) and taking only the fully bound binding mode into account (dashed lines for \( i = 7 \) and \( i = 5 \), respectively) differ by less than a factor of two around the maximum. Since the aim of multivalent ligand design is to improve the minimal dissociation constant by several orders of magnitude, we consider these deviations as negligible.

**Figure S6:** The inverse of the dissociation constant taking all binding modes into account (dotted line) as well as the contributions from all binding modes (dashed lines). The minimal dissociation constant \( K_{\text{min}} \) which corresponds to the maximum of the dotted line, is indicated in each subfigure. a) Heptavalent ligand-receptor pair with \( d_{\text{core}} = 1.5 \text{ nm} \) and \( \omega_{LU} = 0.03 \). b) Pentavalent ligand receptor pair with \( d_{\text{core}} = 0.8 \text{ nm} \) and \( \omega_{LU} = 0.03 \). c) Pentavalent ligand receptor pair with \( d_{\text{core}} = 0.3 \text{ nm} \) and \( \omega_{LU} = 0.03 \).

The balance of the contribution of partially and fully bound ligands to the dissociation constant depends on the value of the angular restriction factor \( \omega_{LU} \), which we determine from a fit of the fully bound contribution to experimental data in the main text. To check whether our fitting procedure is self-consistent, we in Fig.S7a and b show the inverse dissociation constant of a heptavalent ligand-receptor pair for different angular restriction factors \( \omega_{LU} \). For an angular restriction factor \( \omega_{LU} = 0.003 \) in Fig.S7a, the fully bound ligand contribution becomes even more dominant compared to the result for \( \omega_{LU} = 0.03 \) in Fig. S6a, whereas for an angular restriction factor \( \omega_{LU} = 0.3 \) in Fig. S7b, the contribution of partially bound ligands becomes dominant. In Fig. S7c the experimental dissociation constant of the heptavalent anthrax ligand is compared with the theoretical model including all binding modes for \( \omega_{LU} = 0.003, 0.03 \) and 0.3 (the model results correspond to the inverse of the data shown as dotted lines in Fig. S7a, Fig. S6a and Fig. S7b). We also show a curve for the fit value of \( \omega_{LU} = 0.04 \) including all binding modes (continuous line in Fig. S7c). The value \( \omega_{LU} = 0.04 \) is only slightly larger than the value \( \omega_{LU} = 0.03 \) we obtained by a fit of the model that only considers fully bound ligands. We conclude that our fitting procedure for \( \omega_{LU} \) that uses the simplified model that only accounts for fully bound ligands is consistent and gives very similar results as the general model that includes all binding modes and which necessarily needs a numerically demanding Monte Carlo evaluation of the configurational integrals.
Figure S7: Impact of partially bound ligands and of the angular restriction factor $\omega_{LU}$ on the dissociation constant of a heptavalent ligand-receptor pair with $d_{core}=1.5\text{nm}$. a) and b) The inverse of the dissociation constants taking all binding modes into account (dotted line) as well as the contributions from the $i$-th binding modes (dashed lines) for a) $\omega_{LU}=0.003$ and for b) $\omega_{LU}=0.3$. The minimal dissociation constant $K^\text{min}$, which correspond to the maximum of the dotted line, is indicated in each subfigure. c) Comparison of the experimental dissociation constant of the heptavalent anthrax ligand with the theoretical model including all binding modes for different values of $\omega_{LU}$.

Variation of the Core Size for Fixed Linker Length

In Fig.S8a the dissociation constant $K_5$ of a pentavalent ligand is shown in dependence of the core size $d_{core}$ for three fixed linker lengths, corresponding to PEG linkers with $N=10$, 20 or 30 monomers. The receptor size ($d_{rec}=2.5\text{nm}$, indicated by a dotted vertical line) and binding strength ($K_1=5\text{mM}$) are representative for cholera toxin. For comparison the equivalent monovalent dissociation constant $K_1/5=1\text{mM}$ is shown as a black, horizontal line. The results shown in Fig.S8 are based on Eq.S22-S26 and are thus obtained on the same level of approximation as all model results shown in the main text.

Fig.S8a shows that the lowest dissociation constant is achieved with the shortest linker. The dissociation constant of the ligands with longer linkers is two to three orders of magnitude larger, but exhibits a broad plateau and is hence very robust towards variations of the core size. If the core size is larger than the receptor size, $d_{core}>d_{rec}$, the dissociation constant $K_5$ continuously increases with increasing core size, i.e. the better the match between core size and receptor size, the stronger becomes the binding. In contrast, if the core size is smaller than the receptor size, $d_{core}<d_{rec}$, the dissociation constant $K_5$ can increase or decrease with increasing core size depending on the linker length $r_{ete}$. To explain this behavior, we show in Fig.S8b-d all three terms that contribute to the effective concentration in Eq.S22 and thereby contribute to the dissociation constant (Eq.3 in the main text). The angular space of the core, $\Omega_c$, as well as the number of permutations, $\Pi$, decrease with increasing core size, while the conformational weight $Q_c$ exhibits a maximum if core size and receptor size are equal. Whether $\Omega_c$ and $\Pi$ or $Q_c$ dominate the effective concentration depends on the linker length.
Figure S8: a) Dissociation constant $K_5$ of a pentavalent ligand in dependence of the core size $d_{\text{core}}$ for three linker lengths, corresponding to PEG linkers with $N=10$, 20 or 30 monomers. The receptor size $d_{\text{rec}}=2.5\text{nm}$ is indicated as a dotted vertical line. b) Angular space of the ligand core $\Omega_c$ according to Eq.S23, c) number of permutations $\Pi$ according to Eq.S26 and d) conformational weight $Q_c$ according to Eq.S25 in dependence of the core size.

References


