Influence of length and flexibility of spacers on the binding affinity of divalent ligands

Susanne Liese and Roland R. Netz[*]

Full Research Paper

Address:
Fachbereich für Physik, Freie Universität Berlin, 14195 Berlin, Germany

Email:
Roland R. Netz[*] - rnetz@physik.fu-berlin.de

* Corresponding author

Abstract

We present a quantitative model for the binding of divalent ligand–receptor systems. We study the influence of length and flexibility of the spacers on the overall binding affinity and derive general rules for the optimal ligand design. To this end, we first compare different polymeric models and determine the probability to simultaneously bind to two neighboring receptor binding pockets. In a second step the binding affinity of divalent ligands in terms of the IC_{50} value is derived. We find that a divalent ligand has the potential to bind more efficiently than its monovalent counterpart only, if the monovalent dissociation constant is lower than a critical value. This critical monovalent dissociation constant depends on the ligand-spacer length and flexibility as well as on the size of the receptor. Regarding the optimal ligand-spacer length and flexibility, we find that the average spacer length should be equal or slightly smaller than the distance between the receptor binding pockets and that the end-to-end spacer length fluctuations should be in the same range as the size of a receptor binding pocket.

Introduction

Multivalency is a common design principle in biological systems. The simultaneous binding of several, relatively weakly binding partners is a widely used strategy to strengthen the overall binding affinity [1-3]. Multivalency is believed to play an important role in evolutionary processes, since the collective interaction of several rather simple ligands makes the development of more complex binding partners with a higher binding affinity unnecessary [2]. Also in drug design, the synthesis of artificial multivalent ligands is a promising route to increase the binding affinity or to reduce the amount of substance required for treatment [4-7].

The term multivalency is used for systems that consist of several identical binding partners. Thereby, the larger binding partner, for example a protein, is commonly denoted as receptor, whereas the smaller binding partner, for instance an
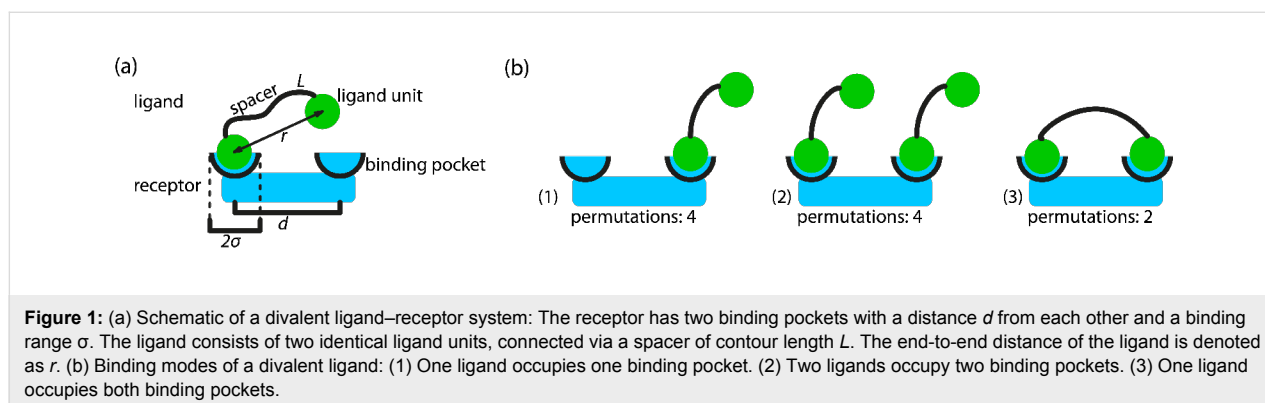
enzyme or a single molecule, is denoted as ligand. The binding strength of a multivalent structure significantly depends on details of the presentation of ligands and receptors [1]. Each multivalent ligand consists of several monovalent ligands that are connected via a scaffold. The binding affinity of such a multivalent ligand is determined by the interplay between gain in binding energy and loss of entropy associated with conformational degrees of freedom. The more flexible the scaffold is, the better it can adapt to the geometry of the receptor, but the more pronounced on the other hand is the entropy penalty. This simple, qualitative argument shows that the careful choice of the ligand scaffold is essential, in order to benefit from multivalent enhancement. It is therefore desirable to derive a model that allows one to predict the binding affinity of a given ligand-scaffold construct. Several theoretical studies have been dedicated to study the interaction between multi- and polyvalent ligands with receptors arranged on planar surfaces [8-13]. The overwhelming variety of multivalent ligand architectures that range from small divalent ligands to densely packed nanoparticles, led to different approaches to describe multivalency, depending on the size and valency of the system. Several studies aimed to treat ligand–receptor systems with different structures and valencies in the framework of a generalized theory [14,15].

The smallest multivalent system consists of a divalent ligand that interacts with a divalent receptor. Despite its seeming simplicity, the rational design of divalent ligands is still challenging [16-19]. In this paper we examine a general model for a divalent receptor–ligand system in order to estimate the binding affinity from the dissociation constant of the monovalent ligand and the length and flexibility of the ligand spacer.

Figure 1a schematically depicts a divalent ligand–receptor system. The receptor possesses two binding pockets at a distance d from each other. A binding range of σ characterizes each binding pocket. The divalent ligand consists of two ligand units that are connected via a spacer. The contour length of the spacer is denoted as L. There are three different modes in which

a divalent ligand can bind to a divalent receptor. Each of these binding modes has a different number of realization possibilities as summarized in Figure 1b: (1) One binding pocket is occupied by one ligand. (2) Two binding pockets are occupied by two ligands. (3) Two binding pockets are occupied by one ligand. The binding affinity in the latter case is strongly influenced by the conformational linker properties, which can be conveniently discussed in terms of the effective concentration. The effective concentration describes the local concentration of one ligand unit close to one binding pocket, if the other ligand unit is assumed to be bound to the other binding pocket. The effective concentration thus corresponds to the probability that the spacer extends to an end-to-end distance that is equal to d, if spacer–receptor interactions are neglected [20]. In the first section different models for the effective concentration are discussed, with particular focus on the influence of the spacer stiffness and the binding range σ.

For each binding mode depicted in Figure 1b the following dissociation constants are derived: (1) The dissociation constant is equal to the dissociation constant of the monovalent ligand, K_{mono}, multiplied by a factor of $1/\alpha$, which accounts for the reduced degrees of freedom of the spacer, since it cannot penetrate the receptor. The parameter α can adopt value between 0 and 1. In the limiting case, in which the spacer sterically inhibits the ligand unit from binding to the receptor, α becomes 0. In the hypothetical case, in which the conformational degrees of freedom of the spacer do not reduce at all when binding to a receptor, the parameter α becomes 1. (2) Each ligand contributes with a factor of K_{mono}/α to the dissociation constant. (3) The dissociation constant consists of the monovalent dissociation constant for each ligand times the probability that the spacer bridges the two binding pockets. A detailed derivation of the dissociation constants is presented in Supporting Information File 1. Furthermore, Figure 1b summarizes the combinatorial factors for each binding mode that count the number of equivalent permutations. We regard the divalent ligands as distinguishable, we note in passing that this could reflect poly-

Figure 1: (a) Schematic of a divalent ligand–receptor system: The receptor has two binding pockets with a distance d from each other and a binding range σ. The ligand consists of two identical ligand units, connected via a spacer of contour length L. The end-to-end distance of the ligand is denoted as r. (b) Binding modes of a divalent ligand: (1) One ligand occupies one binding pocket. (2) Two ligands occupy two binding pockets. (3) One ligand occupies both binding pockets.

meric spacers that exhibit chemical asymmetry. Our final results do not depend on whether we assume indistinguishable ligand units or not.

Results and Discussion
Effective concentration – wormlike-chain model

Samuel and Sinha [21] developed an exact method to describe the conformational statistics of wormlike chains for the whole range from short to long polymers. Their model is applied here to determine the effective concentration C_{eff}, which is equivalent to the end-to-end distance probability distribution, with the normalization $4\pi \int_0^\infty dr\, r^2 C_{eff}(r) = 1$. An example is shown in Figure 2. The length of the fully extended spacer L is set to 5 nm. The effective concentration, i.e., the probability that a spacer of given length and stiffness extends to a certain end-to-end-distance d, is shown for different persistence lengths l_p. The flexible spacer ($l_p = 1$ nm) exhibits a maximum at $d = 0$. Furthermore, the distribution is very broad, indicating that a flexible spacer can easily bridge two binding pockets, even if the spacer length does not exactly match the inter binding pocket distance d. For a slightly stiffer spacer ($l_p = 1.3$ nm), C_{eff} is even broader, but the maximum of C_{eff} is reduced by a factor of about one half and the distribution shows a plateau between $d = 0$ nm and $d = 3$ nm. For stiff spacers ($l_p = 5$ nm and $l_p = 10$ nm), C_{eff} exhibits a narrow peak close to the fully extended state. In the bound state, the ligand units explore the range σ of a receptor binding pocket. Hence, it is useful to consider the effective concentration averaged over the range of both binding pockets. We denote the averaged effective concentration as $\tilde{C}_{eff}$ with

$$\tilde{C}_{eff}(d) = \frac{\int_{V_{bp}} d\mathbf{r}_1 \int_{V_{bp}} d\mathbf{r}_2\, C_{eff}\left(|\mathbf{r}_2 - \mathbf{r}_1|\right)}{\int_{V_{bp}} d\mathbf{r}_1 \int_{V_{bp}} d\mathbf{r}_2}, \qquad (1)$$

with V_{bp} the volume of one binding pocket, $\mathbf{r}_1$ and $\mathbf{r}_2$ the positions within the first and second binding pocket. We introduce the connecting vector $\mathbf{r} = |\mathbf{r}_1 - \mathbf{r}_2|$ and express $\mathbf{r}$ in spherical coordinates:

$$\tilde{C}_{eff}(d) = \frac{\int_{V_{bp}} d\mathbf{r}_1 \int_0^{2\pi} d\varphi \int_{d-2\sigma}^{d+2\sigma} dr\, r^2 \int_0^{\theta(r)} d\theta' \sin(\theta')\, C_{eff}(r)}{\int_{V_{bp}} d\mathbf{r}_1 \int_0^{2\pi} d\varphi \int_{d-2\sigma}^{d+2\sigma} dr\, r^2 \int_0^{\theta(r)} d\theta' \sin(\theta')}, \quad (2)$$

with r the distance between the two ligand units, θ the angle between $\mathbf{r}$ and the connecting vector of the binding pocket midpoints and φ an angle that describes the rotation around the connecting vector of the binding pocket midpoints. Since the

range of the binding pocket σ is assumed to be much smaller than the distance between the binding pockets d, we conclude that the integrals in Equation 2 approximately factorize. Furthermore, the size of the binding pocket limits the range over which the angle θ can vary. In the range, where r varies between $d - \sigma$ and $d + \sigma$, the angle θ can adopt a maximum value of $\arctan(\sigma/r)$. The upper limit for the integration over θ then reads

$$\frac{\int_0^{\arctan\left(\frac{\sigma}{r}\right)} d\theta' \sin(\theta')}{2} \approx \frac{\left(\frac{\sigma}{r}\right)^2}{2}, \text{ valid for } \sigma \ll r.$$

The integration over $\mathbf{r}$ can now be described by variations of r in the range from $d - \sigma$ and $d + \sigma$. With these approximations, Equation 2 can be written as an effective average over one dimension:
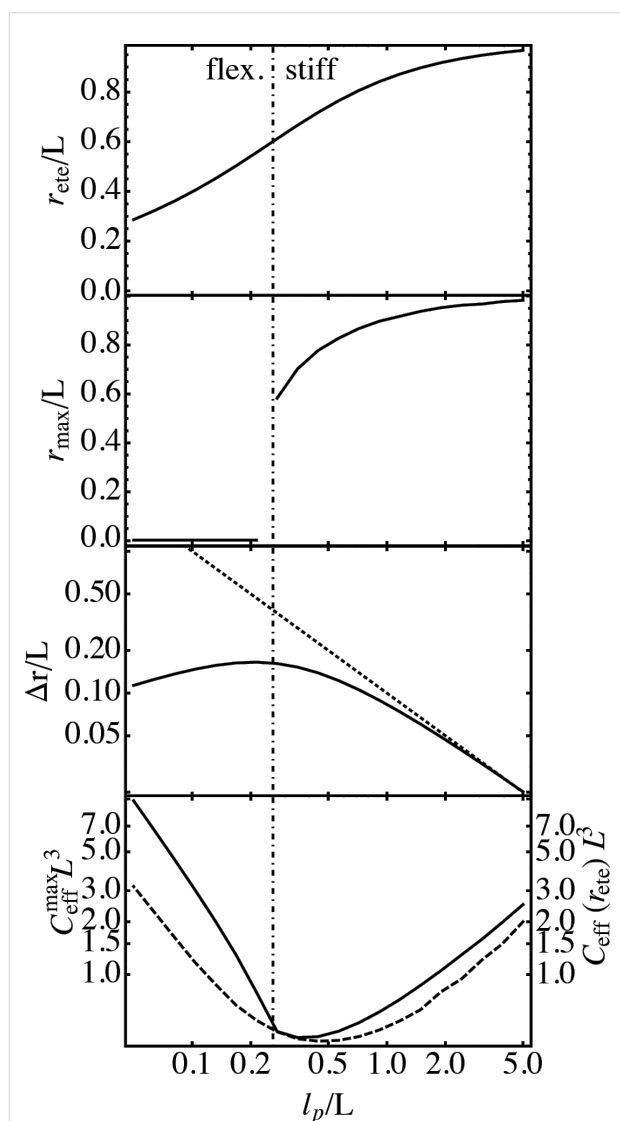
$$\tilde{C}_{eff}(d) = \frac{\int_{d-\sigma}^{d+\sigma} dr\, C_{eff}(r)}{2\sigma}. \qquad (3)$$

In Figure 2, the averaged effective concentration is shown as green, dashed lines, with $\sigma = 0.25$ nm. A flexible spacer can easily extend to all positions within the binding pockets. Hence, one cannot observe any significant difference between $\tilde{C}_{eff}$ and C_{eff}. In contrast, a very stiff spacer cannot explore the whole binding pocket. Therefore, the averaged effective concentration is reduced and slightly broadened around its maximum, as can be seen best in Figure 2 for $l_p = 10$ nm.
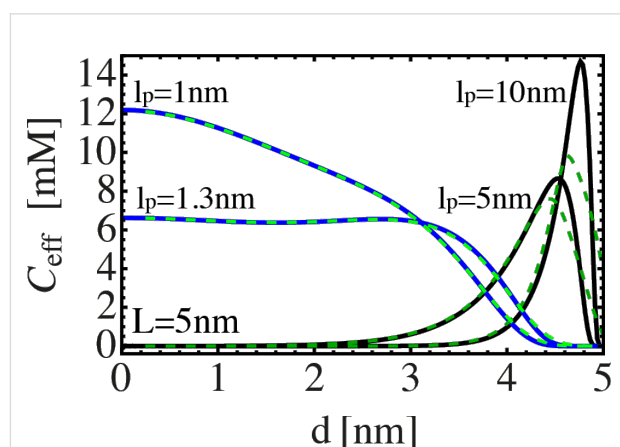
Figure 2: Effective concentration Ceff of spacers with a contour length of $L = 5$ nm as a function of the distance between the binding pockets. The effective concentration is shown for different spacer stiffness, in terms of different persistence lengths between $l_p = 1$–10 nm (continuous lines). The effective concentration $\tilde{C}_{eff}$, averaged over a binding pocket range $\sigma = 0.25$ nm, is shown as green, dashed lines.

Figure 3 summarizes the averaged end-to-end distance r_{ete}, the end-to-end distance that corresponds to a maximum in C_{eff}, r_{max}, the variance of the end-to-end distance distribution Δr, the maximum of the effective concentration C_{eff}^{max} and the effective concentration at r_{ete}, $C_{eff}(r_{ete})$, for different persistence lengths. The influence of the binding range σ is neglected here. The average end-to-end distance r_{ete} increases monotonically with increasing persistence length and approaches the contour length L for very stiff spacers. All other quantities reveal a clear-cut

difference between the flexible and stiff limits. The classification "flexible" and "stiff" is, of course, to some degree arbitrary. We here apply a definition that is based on the discontinuity in r_{max}, which is the most prominent feature in the chain observables. In the following, spacers with a persistence length smaller than $0.26L$ are called flexible and spacers with a persistence length larger than $0.26L$ are called stiff. The variance Δr exhibits a maximum around $l_p = 0.26L$, for stiffer spacers Δr reduces rapidly. As can be seen in Figure 3, the variance Δr depends on the persistence length as $\Delta r = 0.1L^2/l_p$ (dotted line) for stiff spacers. Mac Kintosh et al. found the same scaling for the fluctuations of semiflexible polymers [22]. The maximum of the effective concentration C_{eff}^{max} (continuous line) as well as the effective concentration at r_{ete}, $C_{eff}(r_{ete})$, (dashed line) are minimal in the same region where Δr is maximal. Since for a stiff spacer r_{max} and r_{ete} are both close to L, C_{eff}^{max} and $C_{eff}(r_{ete})$ exhibit only small deviations from each other. For flexible spacers on the other hand, $C_{eff}(r_{ete})$ can be much smaller than the maximal effective concentration. The results presented here show that neither the persistence length nor the contour length alone are sufficient to describe the behavior of the effective concentration, rather the ratio between persistence length and contour length, l_p/L, characterizes the conformational behavior. Note that for a typical receptor distance of $d = 5$ nm, DNA molecules with $l_p = 53$ nm are characterized by a ratio $l_p/L \approx 10$ and thus correspond to the very stiff limit. Polyethylene glycol (PEG) with a persistence length of about $l_p = 0.38$ nm on the other hand is characterized by a ratio smaller than $l_p/L = 0.08$ and thus correspond to the flexible limit [23].

Effective concentration – harmonic spring and Gaussian chain approximation

The wormlike-chain model requires complex numerical analysis for the calculation of conformational chain properties. In a simplified model the spacer statistics can be described as a harmonic spring or a Gaussian chain with suitably chosen parameters. The advantage of this model is that the effective concentration can be derived in closed form. Furthermore, we show that despite its simplified assumptions the model accurately reproduces the effective concentration $C_{eff}(r_{ete})$ for flexible as well as for stiff spacers.

Stiff spacer – harmonic spring approximation

A stiff spacer is on average extended to almost its full length. The fluctuations around its most probable end-to-end distance r_0 are assumed to be much smaller than the contour length L. We approximate the free energy F, similar to a harmonic spring, as

Figure 3: Average end-to-end distance, r_{ete}, end-to-end-distance where the effective concentration C_{eff} exhibits a maximum, r_{max}, variance of the end-to-end distance distribution, Δr, maximum of the effective concentration, C_{eff}^{max} (continuous line), and effective concentration at r_{ete}, $C_{eff}(r_{ete})$ (dashed line), in dependence of the persistence length l_p. All lengths are measured in units of the spacer contour length L. Spacers with a persistence length $l_p < 0.26L$ are called flexible. Spacers with a persistence length $l_p > 0.26L$ are called stiff. For stiff spacers the relation between $\Delta r/L$ and the persistence length is well described by $\Delta r/L = 0.1L/l_p$ (dotted line).

$$F(d) = \frac{k}{2}(d - r_0)^2 , \qquad (4)$$

with k the effective spring constant and d the end-to-end distance. The effective concentration $C_{\text{eff}}(d)$, i.e., the normalized probability to extend the spacer to a certain end-to-end distance d, reads

$$
\begin{aligned}
C_{\text{eff}}(d) &= \frac{e^{-F(d)/k_{\text{B}}T}}{4\pi\int_0^\infty dr\, r^2 e^{-F(r)/k_{\text{B}}T}} \\
&\approx \frac{e^{-F(d)/k_{\text{B}}T}}{4\pi\int_{-\infty}^\infty dr\, r^2 e^{-F(r)/k_{\text{B}}T}} \\
&= \frac{1}{2}\left(\frac{k}{2\pi k_{\text{B}}T}\right)^{3/2}\frac{e^{-F(d)/k_{\text{B}}T}}{1+kr_0^2/k_{\text{B}}T}.
\end{aligned}
\tag{5}
$$

The averaged effective concentration $\tilde{C}_{\text{eff}}$ as defined in Equation 3 then becomes:

$$
\begin{aligned}
\tilde{C}_{\text{eff}}(d) &= \frac{k/k_{\text{B}}T}{16\pi\sigma\left(1+kr_0^2/k_{\text{B}}T\right)} \\
&\times\left(\text{erf}\left[\sqrt{\frac{k}{2k_{\text{B}}T}}\left(r_0-d+\sigma\right)\right]\right. \\
&\left. -\text{erf}\left[\sqrt{\frac{k}{2k_{\text{B}}T}}\left(r_0-d-\sigma\right)\right]\right).
\end{aligned}
\tag{6}
$$

In order to express the effective concentration in term of the experimentally more relevant average end-to-end distance r_{ete} and the variance Δr, we first have to determine the relation between r_{ete} and Δr on the one side and k and r_0 on the other side.

From the free energy F in Equation 4 the average end-to-end distance r_{ete} and the variance Δr are obtained as:

$$
r_{\text{ete}} = \frac{\int_0^\infty dr\, r^3 e^{-F(r)/k_{\text{B}}T}}{\int_0^\infty dr\, r^2 e^{-F(r)/k_{\text{B}}T}} \approx \frac{\int_{-\infty}^\infty dr\, r^3 e^{-F(r)/k_{\text{B}}T}}{\int_{-\infty}^\infty dr\, r^2 e^{-F(r)/k_{\text{B}}T}}.
\tag{7}
$$

$$
\rightarrow r_{\text{ete}} \approx r_0 \frac{3+kr_0^2/k_{\text{B}}T}{1+kr_0^2/k_{\text{B}}T} \quad \rightarrow r_{\text{ete}} \approx r_0;\text{if}\frac{kr_0^2}{k_{\text{B}}T}\gg 1
\tag{8}
$$

$$
\left\langle r_{\text{ete}}^2\right\rangle = \frac{\int_0^\infty dr\, r^4 e^{-F(r)/k_{\text{B}}T}}{\int_0^\infty dr\, r^2 e^{-F(r)/k_{\text{B}}T}} \approx \frac{\int_{-\infty}^\infty dr\, r^4 e^{-F(r)/k_{\text{B}}T}}{\int_{-\infty}^\infty dr\, r^2 e^{-F(r)/k_{\text{B}}T}}
\tag{9}
$$

$$
\rightarrow \left\langle r_{\text{ete}}^2\right\rangle \approx r_0^2 \frac{3k_{\text{B}}T/kr_0^2+6+kr_0^2/k_{\text{B}}T}{1+kr_0^2/k_{\text{B}}T}.
\tag{10}
$$

Note that according to our notation, the average end-to-end distance r_{ete} is not equivalent to the root mean squared end-to-end distance $\sqrt{\left\langle r_{\text{ete}}^2\right\rangle}$. The variance Δr hence reads:

$$
\Delta r = \sqrt{\left\langle r_{\text{ete}}^2\right\rangle - \left(r_{\text{ete}}\right)^2}
\tag{11}
$$

$$
\begin{aligned}
&\rightarrow \Delta r \approx \sqrt{\frac{k_{\text{B}}T}{k}}\sqrt{\frac{3+(kr_0^2/k_{\text{B}}T)^2}{(1+kr_0^2/k_{\text{B}}T)^2}} \\
&\rightarrow \Delta r \approx \sqrt{\frac{k_{\text{B}}T}{k}};\text{if}\frac{kr_0^2}{k_{\text{B}}T}\gg 1.
\end{aligned}
\tag{12}
$$

Using Equation 6 and the results for Δr and r_{ete} in terms of the model parameters k and r_0 in the stiff spacer limit $(kr_0^2/k_{\text{B}}T\gg 1)$, the averaged effective concentration reads:

$$
\tilde{C}_{\text{eff}}(d) = \frac{1}{8\pi\sigma r_{\text{ete}}^2}\left(\text{erf}\left[\frac{r_{\text{ete}}-d+\sigma}{\sqrt{2}\Delta r}\right]-\text{erf}\left[\frac{r_{\text{ete}}-d-\sigma}{\sqrt{2}\Delta r}\right]\right).
\tag{13}
$$

For a fixed distance d that has to be spanned by the ligand, the effective concentration becomes maximal for $r_{\text{ete}}=d$ and we obtain, for this optimized spacer length, the result:

$$
\begin{aligned}
&\tilde{C}_{\text{eff}}(d) = \frac{1}{8\pi\sigma d^2}\text{erf}\left[\frac{\sigma}{\sqrt{2}\Delta r}\right], \\
&\text{with erf}\left[x\right]-\text{erf}\left[-x\right]=2\text{erf}[x].
\end{aligned}
\tag{14}
$$

Furthermore, we can differentiate between two cases: 1) the chain fluctuations are smaller than the binding range ($\Delta r\ll\sigma$) and 2) the chain fluctuations are larger than the binding range ($\Delta r\gg\sigma$), leading to

$$
\Delta r\ll\sigma \quad\rightarrow\quad \tilde{C}_{\text{eff}}^{\text{max}}(d)\approx\frac{1}{8\pi d^2\sigma}
\tag{15}
$$

$$
\Delta r\gg\sigma\rightarrow \tilde{C}_{\text{eff}}^{\text{max}}(d)\approx\frac{1}{2\left(2\pi\right)^{3/2}d^2\Delta r}.
\tag{16}
$$

We see that in both limits, the maximal effective concentration decreases quadratically with the distance d. More importantly, increasing the stiffness of the spacer (decreasing Δr) increases the effective concentration, but only until the variance Δr

becomes of the same order as the binding range σ. For even stiffer spacers the effective concentration stagnates, as can be seen in Equation 15. We conclude that it is not advantageous to increase the spacer stiffness beyond the situation where the end-to-end distance variance Δr becomes smaller than the receptor binding range σ. To compare this model with the wormlike-chain model Equation 16 is rewritten as:

$$\tilde{C}_{\text{eff}}^{\max}\left(d\right)d^3 \approx \frac{1}{2\left(2\pi\right)^{3/2}\Delta r/d}. \tag{17}$$

As can be seen in Figure 4a Equation 17 describes the behavior of stiff wormlike chains very well.

Flexible spacer – Gaussian-chain approximation

The effective concentration of flexible polymers is often modeled by a Gaussian chain [11,20,24] with the free energy:

$$F\left(d\right) = \frac{3}{2}\frac{d^2}{\left\langle r_{\text{ete}}^2\right\rangle}k_{\text{B}}T, \tag{18}$$

using the mean squared end-to-end distance $\left\langle r_{\text{ete}}^2\right\rangle$. The end-to-end distance r_{ete} and the variance Δr can be expressed in terms of the mean squared end-to-end distance:

$$r_{\text{ete}} = \frac{\int_0^\infty dr\, r^3 e^{-F(r)/k_{\text{B}}T}}{\int_0^\infty dr\, r^2 e^{-F(r)/k_{\text{B}}T}} = \sqrt{\frac{8}{3\pi}\left\langle r_{\text{ete}}^2\right\rangle}, \tag{19}$$

$$\Delta r = \sqrt{\left\langle r_{\text{ete}}^2\right\rangle - \left(r_{\text{ete}}\right)^2} = \sqrt{\left(1-\frac{8}{3\pi}\right)\left\langle r_{\text{ete}}^2\right\rangle}. \tag{20}$$

As a consequence the end-to-end distance r_{ete} and the variance Δr are related as

$$\Delta r = r_{\text{ete}}\sqrt{\frac{3\pi}{8}-1}. \tag{21}$$

Furthermore, the mean squared end-to-end distance can be written as

$$\left\langle r_{\text{ete}}^2\right\rangle = b^2 N, \tag{22}$$

with b being the Kuhn length of one chain segment and N the number of segments.

We here present the effective concentration as a function of d and r_{ete}.

$$\begin{aligned} C_{\text{eff}}\left(d\right) &= \frac{e^{-F(d)/k_{\text{B}}T}}{4\pi\int_0^\infty dr\, r^2 e^{-F(r)/k_{\text{B}}T}} \\ &= \frac{1}{r_{\text{ete}}^3}\left(\frac{4}{\pi^2}\right)^{3/2}\exp\left[-\frac{4}{\pi}\left(\frac{d}{r_{\text{ete}}}\right)^2\right]. \end{aligned} \tag{23}$$

Using Equations 19–22, r_{ete} can as well be substituted by $\left\langle r_{\text{ete}}^2\right\rangle$, Δr or N.

Note that the effective concentration of a flexible spacer with fixed contour length L is maximal at a distance $d = 0$, as shown in Figure 2. In contrast, for a given distance d the effective concentration becomes maximal at $r_{\text{ete}} = \sqrt{8/(3\pi)}d$. In other
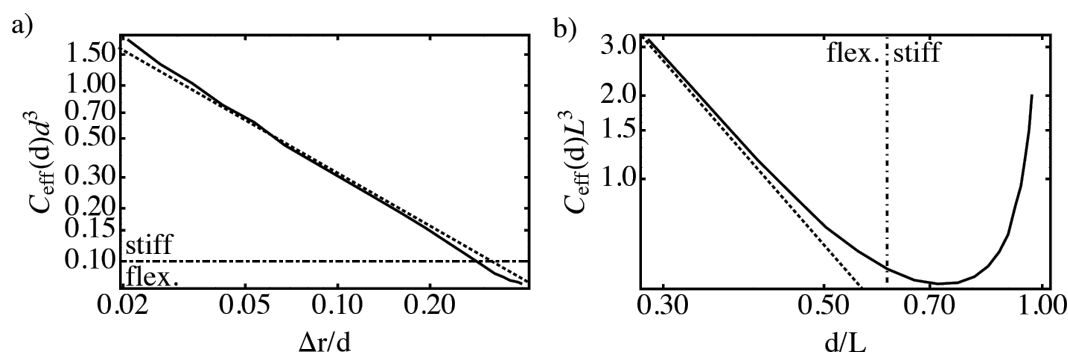
Figure 4: Effective concentration for the optimized average end-to-end distance $r_{\text{ete}}=d$ for the wormlike chain model (continuous line) and the harmonic spring model Equation 17 (dotted line, subfigure a) as well as the Gaussian-chain model Equation 25 (dotted line, subfigure b). In the calculation, we vary the ratio between persistence length and contour length l_{p}/L, which results in different ratios $\Delta r/d$ and d/L, respectively. (a) Stiff spacers are well approximated by Equation 17. (b) Flexible spacers are well approximated by Equation 25.

words, the average end-to-end distance of an optimized flexible spacer is smaller than the distance between the binding pockets by a factor of $\sqrt{8/(3\pi)}$:

$$C_{\text{eff}}^{\max}\left(d\right)=\left(\frac{3}{2\pi d^2}\right)^{3/2}e^{-3/2}\ . \tag{24}$$

Since we consider the fluctuations of a flexible chain much larger than the range of the binding pocket, we neglect the influence of σ on the effective concentration. In order to compare the behavior of a Gaussian chain with the results for a flexible wormlike chain, Equation 24 is rewritten as:

$$C_{\text{eff}}\left(d\right)L^3=\left(\frac{L}{d}\right)^3\left(\frac{4}{\pi}\right)^{3/2}e^{-4/\pi}\ . \tag{25}$$

In Figure 4b, Equation 25 is shown together with the numerical results from the wormlike chain model obtained in the previous section. The two models show good agreement in the flexible limit, as expected.

Conformational degrees of freedom of a tethered spacer

If one ligand unit is bound to one of the binding pockets, the conformational degrees of freedom of the spacer are reduced, since it cannot penetrate the receptor surface. We quantify this reduction by the parameter α, which describes the ratio between the partition function of a tethered and a free spacer. The value of α depends on the shape of the receptor and the flexibility of the spacer. To estimate the typical magnitude of α we consider as limiting cases a stiff rod as well as a flexible Gaussian chain tethered to a planar surface.

Stiff spacer

For a stiff rod attached with one end to a planar surface, the parameter α becomes α = 1/2, since the rod can only explore one half space.

Flexible spacer

As a second example we discuss a Gaussian chain. Equivalently to Equation 23 the normalized probability that a Gaussian chain consisting of N segments extends to an end-to-end distance r with b being the length of one segment reads in free space:

$$P\left(r,N\right)=\left(\frac{3}{2\pi Nb^2}\right)^{3/2}\exp\left[-\frac{3}{2}\frac{r^2}{Nb^2}\right]. \tag{26}$$

We now assume that one end of the chain is attached to the surface. Similar to the considerations made for a stiff rod, we approximate the probability that the first segment does not penetrate the surface by a factor 1/2. The probability distribution for the remaining $N-1$ segments then reads:

$$\begin{aligned}P'\left(\rho,z,N\right)=&\frac{1}{2}\left(\frac{3}{2\pi(N-1)b^2}\right)^{3/2}\exp\left[-\frac{3}{2}\frac{\rho^2}{(N-1)b^2}\right]\\&\times\left(\exp\left[-\frac{3}{2}\frac{(z-b)^2}{(N-1)b^2}\right]+\exp\left[-\frac{3}{2}\frac{(z+b)^2}{(N-1)b^2}\right]\right),\end{aligned} \tag{27}$$

with ρ the component of the end-to-end vector parallel to the surface and z the height above the surface. The last term in Equation 27 ensures that the chain does not penetrate the surface ($P'(\rho,z=0,N)=0$). To obtain the parameter α, P' has to be integrated over one half space:

$$\begin{aligned}\alpha(N)=&\frac{1}{2}\left(\frac{3}{2\pi(N-1)b^2}\right)^{1/2}\\&\times\int_0^\infty dz\left(\exp\left[-\frac{3}{2}\frac{(z-b)^2}{(N-1)b^2}\right]-\exp\left[-\frac{3}{2}\frac{(z+b)^2}{(N-1)b^2}\right]\right).\end{aligned} \tag{28}$$

In the limit of a long chain ($N \gg 1$), Equation 28 can be approximated as:

$$\begin{aligned}\alpha\left(N\right)=&\frac{1}{2}\left(\frac{3}{2\pi Nb^2}\right)^{\frac{1}{2}}\frac{6}{Nb}\exp\left[-\frac{3}{2N}\right]\int_0^\infty dz\,z\exp\left[-\frac{3}{2}\frac{z^2}{Nb^2}\right]\\=&\sqrt{\frac{3}{2\pi N}}\exp\left[-\frac{3}{2N}\right].\end{aligned} \tag{29}$$

A PEG spacer with $b = 0.38$ nm requires 30–800 segments to adopt an average end-to-end distance of 2 to 10 nm. In this range α varies between 0.02 and 0.13.

Binding affinity

With the effective concentration and a parameterization for the reduction of the conformational degrees of freedom of the spacer at hand, we now can examine the binding affinity of a divalent ligand. A common way to quantify the binding affinity of a multivalent ligand is the so-called IC_{50} value, the ligand (or inhibitor) concentration at half maximal inhibition. In a first step we want to re-derive the relation between the IC_{50} value and the dissociation constant of a monovalent ligand [25,26].

Monovalent ligand

In the reaction $L + R \rightleftharpoons LR$, the dissociation constant K_{mono} of a monovalent ligand interacting with a monovalent receptor is defined as

$$K_{mono} = \frac{[L][R]}{[RL]}, \qquad (30)$$

with [L] and [R] being the concentration of unbound ligands and unbound receptors and [RL] the concentration of bound ligands or equivalently the concentration of bound receptors.

If half of all receptors are occupied, which defines the IC_{50} condition, the other half must be unbound and as a consequence [R] = [RL]. From Equation 30 we see that under IC_{50} conditions the dissociation constant equals the concentration of unbound ligands:

$$K_{mono} = [L]_{50}, \qquad (31)$$

with the index 50 indicating that the IC_{50} condition is fulfilled. In the monovalent case exactly one ligand binds to one receptor. Thus, the concentration of bound ligands under IC_{50} conditions is given by half the total receptor concentration:

$$[RL]_{50} = \frac{1}{2}[R]_0, \qquad (32)$$

with $[R]_0 = [R] + [RL]$ the total receptor concentration. Combining Equation 31 and Equation 32 the IC_{50} value is obtained as [25]:

$$IC_{50} \equiv [L]_{0,50} = [L]_{50} + [RL]_{50} = K_{mono} + \frac{1}{2}[R]_0. \quad (33)$$

In the limit of dilute receptor conditions ($[R]_0 \ll K_{mono}$) the IC_{50} value is a good approximation for the dissociation constant, and we find:

$$IC_{50} \approx K_{mono}, \text{if } [R]_0 \ll K_{mono}. \qquad (34)$$

Divalent ligand

In analogy to the monovalent case, we now derive an expression for the IC_{50} value of a divalent ligand. There are different ways of defining half maximal inhibition for divalent receptors. We first adopt a heuristic definition where half of all receptor binding pockets are occupied by a ligand unit. This definition is most relevant for competitive binding assays, for instance surface plasmon resonance measurements [27], since the

measured signal in a competitive binding assay is related to the number of occupied binding pockets. Later, we also define a situation in which at least one ligand unit is bound to half of all receptors as IC_{50} condition, which mimics non-competitive binding assays, as for instance hemagglutination assays [28]. In non-competitive binding assays the number of bound ligands rather than the number of occupied binding pockets is measured. In general the concentration of occupied binding pockets $[bp]_{occ}$ of divalent receptors reads:

$$\begin{aligned} [bp]_{occ} &= 1 \cdot [RL_1] + 2 \cdot [RL_2] + 2 \cdot [RL_3] \\ &= 1 \cdot 4 \cdot \alpha \frac{[L][R]}{K_{mono}} \\ &\quad + 2 \cdot 4 \cdot \alpha^2 \frac{[L]^2[R]}{K_{mono}^2} \\ &\quad + 2 \cdot 2 \cdot \frac{\tilde{C}_{eff}[L][R]}{K_{mono}^2}, \end{aligned} \qquad (35)$$

with $[RL_n]$ being the concentration of bound ligand–receptor pairs, with n referring to the three binding modes summarized in Figure 1b. Each term on the right hand side of Equation 35 has two prefactors. The first prefactor counts the number of occupied binding pockets per receptor and the second prefactor counts the permutations due to the distinguishability of the ligand units and the receptor binding pockets (see Figure 1b). Note that the number of permutations presented in Figure 1b and Equation 35, are obtained for distinguishable ligand units. For indistinguishable ligand units the number of permutations in each binding mode is reduced. At the same time the dissociation constant of a ligand with indistinguishable ligand units is reduced by the same factor. Hence, the overall concentration of bound ligands does not change. A detailed derivation of the dissociation constants for each binding mode is presented in Supporting Information File 1.

In the same way the total concentration of binding pockets, $[bp]_0$, can be obtained as

$$\begin{aligned} [bp]_0 &= 2 \cdot ([R] + [RL_1] + [RL_2] + [RL_3]) \\ &= 2 \cdot \Big([R] + 4 \cdot \alpha \frac{[L][R]}{K_{mono}} \\ &\quad + 4 \cdot \alpha^2 \frac{[L]^2[R]}{K_{mono}^2} \\ &\quad + 2 \cdot \frac{\tilde{C}_{eff}[L][R]}{K_{mono}^2} \Big). \end{aligned} \qquad (36)$$

In order to discuss also the IC_{50} condition for non-competitive binding assays we derive the concentration of receptors with at least one binding pocket occupied, $[R]_{1bp}$, and the total receptor concentration, $[R]_0$, as

$$[R]_{1bp} = [RL_1] + [RL_2] + [RL_3]$$
$$= \left(4 \cdot \alpha \frac{[L]}{K_{mono}} + 4 \cdot \alpha^2 \frac{[L]^2}{K_{mono}^2} + 2 \cdot \frac{\tilde{C}_{eff}[L]}{K_{mono}^2} \right)[R], \quad (37)$$

$$[R]_0 = [R] + [RL_1] + [RL_2] + [RL_3]$$
$$= \left(1 + 4 \cdot \alpha \frac{[L]}{K_{mono}} + 4 \cdot \alpha^2 \frac{[L]^2}{K_{mono}^2} + 2 \cdot \frac{\tilde{C}_{eff}[L]}{K_{mono}^2} \right)[R]. \quad (38)$$

With Equations 35–38 the IC_{50} condition for competitive and non-competitive binding is expressed as given in Equation 39 and Equation 40.

In analogy to the monovalent case we define the multivalent dissociation constant K_{multi} as the concentration of free ligand under IC_{50} conditions, as defined in Equation 39 and Equation 40.

Equation 41 and Equation 42 show the multivalent dissociation constant K_{multi} in case of competitive binding and non-competitive binding, respectively.

$$K_{multi} \equiv [L]_{50}$$
$$= K_{mono} \frac{1}{4\alpha^2} \left(\sqrt{\left(\frac{\tilde{C}_{eff}}{K_{mono}} \right)^2 + 4\alpha^2} - \frac{\tilde{C}_{eff}}{K_{mono}} \right)$$
$$\approx \begin{cases} \dfrac{1}{2\alpha} K_{mono} & ; \text{if } \dfrac{\tilde{C}_{eff}}{K_{mono}} \ll 1 \\[2ex] \dfrac{K_{mono}^2}{2\tilde{C}_{eff}} & ; \text{if } \dfrac{\tilde{C}_{eff}}{K_{mono}} \gg 1 \end{cases} \quad (41)$$

Competitive and non-competitive binding exhibit the same qualitative behavior for large effective concentrations. We therefore limit the further discussion to competitive binding, as given in Equation 41.

As one would intuitively expect, the multivalent dissociation constant K_{multi} becomes proportional to the monovalent dissociation constant, if the effective concentration is low, i.e., if $\tilde{C}_{eff} \ll K_{mono}$. In contrast, the multivalent dissociation constant

$$\text{competitive binding:} \quad \frac{1}{2} = \frac{[bp]_{occ}}{[bp]_0} = \frac{1 \cdot 4 \cdot \alpha \dfrac{[L]}{K_{mono}} + 2 \cdot 4 \cdot \alpha^2 \dfrac{[L]^2}{K_{mono}^2} + 2 \cdot 2 \cdot \dfrac{\tilde{C}_{eff}[L]}{K_{mono}^2}}{2 \cdot \left(1 + 4 \cdot \alpha \dfrac{[L]}{K_{mono}} + 4 \cdot \alpha^2 \dfrac{[L]^2}{K_{mono}^2} + 2 \cdot \dfrac{\tilde{C}_{eff}[L]}{K_{mono}^2} \right)} \quad (39)$$

$$\text{non-competitive binding:} \quad \frac{1}{2} = \frac{[R]_{1bp}}{[R]_0} = \frac{4 \cdot \alpha \dfrac{[L]}{K_{mono}} + 4 \cdot \alpha^2 \dfrac{[L]^2}{K_{mono}^2} + 2 \cdot \dfrac{\tilde{C}_{eff}[L]}{K_{mono}^2}}{1 + 4 \cdot \alpha \dfrac{[L]}{K_{mono}} + 4 \cdot \alpha^2 \dfrac{[L]^2}{K_{mono}^2} + 2 \cdot \dfrac{\tilde{C}_{eff}[L]}{K_{mono}^2}} \quad (40)$$

$$K_{multi} \equiv [L]_{50}$$
$$= K_{mono} \frac{1}{4\alpha^2} \left(\sqrt{\left(\frac{\tilde{C}_{eff}}{K_{mono}} + 2\alpha \right)^2 + 4\alpha^2} - \left(\frac{\tilde{C}_{eff}}{K_{mono}} + 2\alpha \right) \right)$$
$$\approx \begin{cases} \left(\sqrt{2} - 1 \right) \dfrac{1}{2\alpha} K_{mono} & ; \text{if } \dfrac{\tilde{C}_{eff}}{K_{mono}} \ll 1 \\[2ex] \dfrac{K_{mono}^2}{2\tilde{C}_{eff}} & ; \text{if } \dfrac{\tilde{C}_{eff}}{K_{mono}} \gg 1 \end{cases} \quad (42)$$

decreases, if the dissociation constant of the monovalent ligand is small and if the effective concentration, i.e., the probability to connect two binding pockets, is large.

To determine the total ligand concentration we first have to derive the concentration of bound ligand [L]$_{bound}$ as shown in Equation 43.

Using Equation 38 and 43, a relation between the concentration of bound ligands and the total receptor concentration under IC$_{50}$ conditions is obtained as

$$\left[\mathrm{L}\right]_{\mathrm{bound},50} = \psi\left[\mathrm{R}\right]_0, \text{ with}$$

$$\psi = \frac{\left(\sqrt{\left(\dfrac{\tilde{C}_{\mathrm{eff}}}{2\alpha K_{\mathrm{mono}}}\right)^2 + 1} - \dfrac{\tilde{C}_{\mathrm{eff}}}{2\alpha K_{\mathrm{mono}}}\right)\left(1 - \dfrac{\tilde{C}_{\mathrm{eff}}}{2\alpha K_{\mathrm{mono}}}\right) + \dfrac{3}{2}}{\sqrt{\left(\dfrac{\tilde{C}_{\mathrm{eff}}}{2\alpha K_{\mathrm{mono}}}\right)^2 + 1} - \dfrac{\tilde{C}_{\mathrm{eff}}}{2\alpha K_{\mathrm{mono}}} + 1}, \quad (44)$$

$$IC_{50} = \left[\mathrm{L}\right]_{50} + \left[\mathrm{L}\right]_{\mathrm{bound},50} = K_{\mathrm{multi}} + \psi\left[\mathrm{R}\right]_0, \quad (45)$$

where we note that that ψ is a coefficient that varies between 1 and 5/4. Similar to the results for monovalent receptor–ligand systems in Equation 34, the IC$_{50}$ value becomes equivalent to the multivalent dissociation constant, in the limit of low receptor concentrations, i.e., for $[\mathrm{R}]_0 \ll K_{\mathrm{multi}}$:

$$IC_{50} \approx K_{\mathrm{multi}} \text{ ; if } \left[\mathrm{R}\right]_0 \ll K_{\mathrm{multi}}. \quad (46)$$

To compare monovalent and multivalent ligands we use the relative binding affinity (*RBA*), which we define as

$$RBA = \frac{K_{\mathrm{mono}}}{2K_{\mathrm{multi}}}. \quad (47)$$

Here, the factor 2 accounts for the valency of the ligand and ensures that the concentration of ligand units are compared. The larger the *RBA* the better is the divalent ligand. For *RBA* = 1 the same concentration of mono- and divalent ligand units, taking into account that a divalent ligand consist of two ligand units, is required to occupy half of the receptor binding pockets. For *RBA* < 1 the monovalent ligand binds better than the divalent ligand. In this case the loss in entropy of the spacer is larger

than the gain in binding energy due to the multiple binding of ligand units. Inserting the effective concentration from Equation 13 and Equation 23 into Equation 41 and Equation 47, the *RBA* can be calculated for any given divalent ligand–receptor pair. As an example the *RBA* is depicted for different spacers and different values of K_{mono} in Figure 5. We here assume that the receptor is well described by a large, planar surface. Hence, the parameter α is approximated by 1/2 for stiff spacer and by Equation 29 for flexible spacers. In all cases we consider a divalent receptor with a distance $d = 5$ nm between the binding pockets. Each binding pocket has a binding range $\sigma = 0.1$ nm. In all three subfigures we see that if K_{mono} is too large, i.e., if the monovalent binder is too weak, the *RBA*-
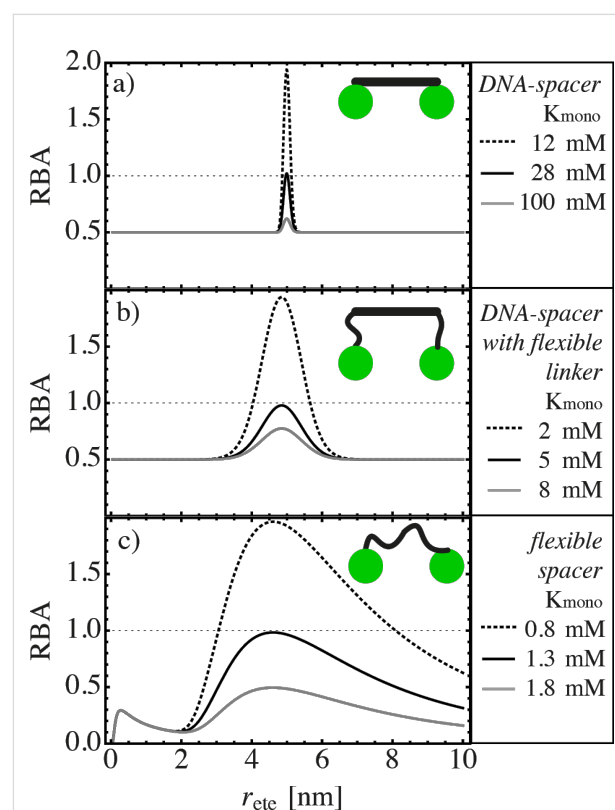
Figure 5: Relative binding affinity (*RBA*) of a divalent ligand in dependence of the end-to-end distance of the spacer r_{ete} from Equation 47. The three different ligand–spacer constructs are schematically depicted in the insets. The binding pockets are separated by $d = 5$ nm. Each binding pocket has a binding range of $\sigma = 0.1$ nm. (a) The ligand units are directly attached to a stiff DNA spacer, characterized by a persistence length $l_p = 53$ nm. (b) The ligand units are attached to a stiff DNA spacer with flexible linker chain, leading to an end-to-end distance fluctuation of $\Delta r = 0.5$ nm. (c) The ligand units are connected via a flexible spacer.

$$\left[\mathrm{L}\right]_{\mathrm{bound}} = 1 \cdot \left[\mathrm{RL}_1\right] + 2 \cdot \left[\mathrm{RL}_2\right] + 1 \cdot \left[\mathrm{RL}_3\right] = \left(1 \cdot 4 \cdot \alpha \frac{\left[\mathrm{L}\right]}{K_{\mathrm{mono}}} + 2 \cdot 4 \cdot \alpha^2 \frac{\left[\mathrm{L}\right]^2}{K_{\mathrm{mono}}^2} + 1 \cdot 2 \cdot \frac{\tilde{C}_{\mathrm{eff}}\left[\mathrm{L}\right]}{K_{\mathrm{mono}}^2}\right)\left[\mathrm{R}\right]. \quad (43)$$

value never reaches 1. In such a situation, using the *RBA*-value as a quantifier, the monovalent ligand binds always better than the divalent ligand. Furthermore, at a certain K_{mono}, which we will further on denote as K_{mono}^*, there is exactly one spacer length, parameterized by r_{ete}, for which monovalent and divalent ligands bind equally well. If K_{mono} is lower than K_{mono}^*, there is a broader range of spacer lengths for which the divalent ligand binds better than the monovalent ligand (*RBA* > 1). In Figure 5a the behavior of a stiff spacer with persistence length l_{p} = 53 nm is depicted, which mimics a DNA spacer to which the ligand units are directly attached. A DNA spacer with a contour length of 5 nm exhibits fluctuations in the range $\Delta r \approx$ 0.05 nm, which is considerably smaller than the binding range σ. As is discussed in the previous section, the maximum and width of the effective concentration and therefore also the maximum and width of the *RBA* are in this case determined by the binding range σ. In Figure 5b we assume a DNA spacer that is decorated with flexible PEG linkers at both ends. The PEG linkers consist of four monomers each. Assuming Gaussian-chain behavior with a segment length of b = 0.38 nm [29], the fluctuations of the PEG linkers and hence the fluctuations of the whole ligand sum up to Δr = 0.5 nm. The shape of the *RBA* now is much broader, showing that the ligand is less affected by a mismatch between spacer length and distance between the binding pockets. Additionally, we obtain K_{mono}^* = 5 mM in Figure 5b which is considerably smaller than K_{mono}^* = 28 mM for the pure DNA spacer in Figure 5a. The same trend is continued in Figure 5c. The more flexible the spacer, the smaller is K_{mono}^*, indicating that flexible spacers are less suitable to improve the binding affinity of weak monovalent binders, even though they are more tolerant with respect to a mismatch between linker length and receptor distance.

To investigate the transition from *RBA* < 1 to *RBA* > 1 further, we determine the critical dissociation constant K_{mono}^* for which the *RBA* is equal to one for the optimized chain length, i.e., for the chain length that maximizes the *RBA* value. Using Equation 41 and Equation 47 it can easily be seen that K_{mono}^* relates to the effective concentration $\tilde{C}_{\mathrm{eff}}(d)$ as

$$K_{\mathrm{mono}}^* = \frac{1}{1-\alpha^2} \tilde{C}_{\mathrm{eff}}(d). \qquad (48)$$

In Figure 6, K_{mono}^* is shown for stiff as well as flexible ligands. The stiff ligand is considered to consist of a DNA spacer to which the ligand units are attached via two PEG linkers. Linker length and binding range are set to be identical to the example presented in Figure 5b. The average end-to-end distance of the DNA spacer is either chosen to be equal to d (black, continuous line), or is chosen to be too short by 0.7 nm, which mimics the length of two base pairs (red, continuous line). Even though the

mismatch between spacer length and binding pocket distance is small, the ligand becomes significantly less efficient.

The flexible ligand is chosen to resemble a PEG spacer. Again, we assume Gaussian-chain behavior with a segment length of b = 0.38 nm. A ligand with optimized spacer length (black, dashed line) does not exhibit a significant difference to a ligand with a spacer that is shortened by two segments (red, dashed line). This shows again that a flexible chain is more tolerant with respect to a distance mismatch between inter-binding pocket distance d and chain length.

If the monovalent dissociation constant is larger than K_{mono}^*, a monovalent ligand always binds better than a divalent ligand. On the other hand, if the monovalent dissociation constant is smaller than K_{mono}^*, a divalent ligand of optimally (or slightly suboptimal) chosen size binds better than a monovalent ligand.

As can be seen in Figure 6, K_{mono}^* depends on the distance between the binding pockets as well as the spacer length and flexibility. In order to approximate an upper limit for K_{mono}^*, the maximum effective concentration (Equation 24 for a flexible spacer and Equation 15 and Equation 16 for a stiff spacer) is substituted into Equation 48:

$$\text{flexible spacer: } K_{\mathrm{mono}}^* \leq \frac{1}{1-\alpha^2}\left(\frac{3}{2\pi d^2}\right)^{3/2} e^{-3/2}, \quad (49)$$

$$\text{stiff spacer: } K_{\mathrm{mono}}^* \leq \begin{cases} \dfrac{1}{1-\alpha^2}\dfrac{1}{8\pi^2 d^2 \sigma} & ; \Delta r \ll \sigma \\[2ex] \dfrac{1}{1-\alpha^2}\dfrac{1}{2(2\pi)^{3/2} d^2 \Delta r} & ; \Delta r \gg \sigma \end{cases} \quad (50)$$

As an example that is relevant for medical applications we want to briefly discuss the interaction between hemagglutinin (HA), a receptor protein on the surface of influenza viruses, and its ligand sialic acid (SA). The dissociation constant between monomeric SA and trimeric HA is known to be 2.5 mM [1]. Furthermore, the crystal structure of HA [30] indicates a distance between neighboring binding pockets in the range of d = 5 nm. Note that HA is a trivalent receptor, which means that additional binding modes as well as different numbers of permutations (see Figure 1b) have to be considered. Nevertheless, since the efficiency of a divalent ligand is mainly influenced by the effective concentration $\tilde{C}_{\mathrm{eff}}$ and the monovalent dissociation constant K_{mono}, rather than by the number of binding modes, we can compare the values for the SA–HA pair

with the results presented in Figure 6. We see that a divalent ligand consisting of two SA units connected via a PEG spacer is expected to bind less efficient than the monovalent SA. In contrast, a stiff DNA spacer can increase the binding affinity of the divalent ligand compared to the monovalent ligand, if its length is optimized.
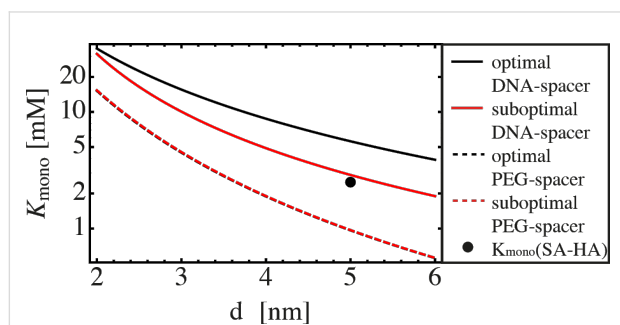
Figure 6: Efficiency diagram: K^*_{mono} is shown for different ligand–spacer constructs. If the monovalent dissociation constant is larger than K^*_{mono}, a monovalent ligand always binds better than a divalent ligand. If, on the other hand, the monovalent dissociation constant is smaller than K^*_{mono}, a divalent ligand of suitably chosen length binds better than its monovalent counterpart. We present K^*_{mono} in dependence of the distance between the binding pockets for a DNA spacer with flexible PEG linkers ($\Delta r = 0.5$ nm). In the optimal case, the spacer length is chosen equal to the distance d (black, continuous line). In the slightly suboptimal case, the spacer length is chosen to be 0.7 nm (two base pairs) shorter than the distance d (red, continuous line). In both cases the binding range is set to $\sigma = 0.1$ nm. We also show K^*_{mono} for a flexible PEG spacer with optimized spacer length (black, dashed line) and a spacer that is two monomers shorter (≈ 0.76 nm) (red, dashed line). The monovalent dissociation constant K^*_{mono} as well as the distance between neighboring binding pockets for a SA–HA pair is indicated by a black point.

Conclusion

In the present work we first examine different polymeric models for the effective concentration. We find that a worm-like-chain model can be well reproduced by a simple harmonic spring model and a Gaussian-chain model with suitable chosen parameters, in the stiff and flexible limits, respectively. We next study the binding between divalent ligand–receptor pairs. We find that multivalency increases the overall binding affinity only, if the monovalent ligand–receptor pair binds strongly enough, i.e.; if the monovalent dissociation constant is smaller than a critical value K^*_{mono}. Approximations for K^*_{mono} for both flexible and stiff ligands are derived in dependence of the distance between the binding pockets and the spacer length and flexibility. For the optimal ligand design, we find that for stiff ligands the average end-to-end distance should be equal to the distance between the binding pockets and the average fluctuations should be of the order, but not smaller, than the binding range. The average end-to-end distance of a flexible ligand on the other side should be smaller by a factor of $\sqrt{8/(3\pi)}$ than the binding pocket distance d.

Supporting Information

Supporting Information File 1

Detailed derivation of the dissociation constants for three different binding modes of a divalent ligand.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-11-90-S1.pdf]

References

1. Mammen, M.; Choi, S.-K.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **1998,** *37,* 2754–2794. doi:10.1002/(SICI)1521-3773(19981102)37:20<2754::AID-ANIE2754>3.0.CO;2-3
2. Kiessling, L. L.; Young, T.; Mortell, K. H. Multivalency in Protein–Carbohydrate Recognition. In *Glycoscience – Chemistry and Chemical Biology;* Fraser-Reid, B. O.; Tatsuka, K.; Thiem, J., Eds.; Springer: Berlin, Germany, 2001; pp 1817–1861. doi:10.1007/978-3-642-56874-9_42
3. Pieters, R. J. *Org. Biomol. Chem.* **2009,** *7,* 2013–2025. doi:10.1039/b901828j
4. Disney, M. D.; Zheng, J.; Swager, T. M.; Seeberger, P. H. *J. Am. Chem. Soc.* **2004,** *126,* 13343–13346. doi:10.1021/ja047936i
5. Wang, J.; Tian, S.; Petros, R. A.; Napier, M. E.; DeSimone, J. M. *J. Am. Chem. Soc.* **2010,** *132,* 11306–11313. doi:10.1021/ja1043177
6. Schaschke, N.; Matschiner, G.; Zettl, F.; Marquardt, U.; Bergner, A.; Bode, W.; Sommerhoff, C. P.; Moroder, L. *Chem. Biol.* **2001,** *8,* 313–327. doi:10.1016/S1074-5521(01)00011-4
7. Vance, D.; Shah, M.; Joshi, A.; Kane, R. S. *Biotechnol. Bioeng.* **2008,** *101,* 429–434. doi:10.1002/bit.22056
8. Martinez-Veracoechea, F. J.; Frenkel, D. *Proc. Natl. Acad. Sci. U. S. A.* **2011,** *108,* 10963–10968. doi:10.1073/pnas.1105351108
9. Hu, J.; Lipowsky, R.; Weikl, T. R. *Proc. Natl. Acad. Sci. U. S. A.* **2013,** *110,* 15283–15288. doi:10.1073/pnas.1305766110
10. Wang, S.; Dormidontova, E. E. *Phys. Rev. Lett.* **2012,** *109,* 238102. doi:10.1103/PhysRevLett.109.238102
11. Diestler, D. J.; Knapp, E. W. *J. Phys. Chem. A* **2010,** *114,* 5287–5304. doi:10.1021/jp100077n
12. Weber, M.; Bujotzek, A.; Haag, R. *J. Chem. Phys.* **2012,** *137,* 054111. doi:10.1063/1.4739501
13. Huskens, J.; Mulder, A.; Auletta, T.; Nijhius, C. A.; Ludden, M. J. W.; Reinhoudt, D. N. *J. Am. Chem. Soc.* **2004,** *126,* 6784–6797. doi:10.1021/ja049085k
14. Varilly, P.; Angioletti-Uberti, S.; Mognetti, B. M.; Frenkel, D. *J. Chem. Phys.* **2012,** *137,* 094108. doi:10.1063/1.4748100
15. Angioletti-Uberti, S.; Varilly, P.; Mognetti, B. M.; Tkachenko, A. V.; Frenkel, D. *J. Chem. Phys.* **2013,** *138,* 021102. doi:10.1063/1.4775806
16. Pertici, F.; de Mol, N. J.; Kemmink, J. M.; Pieters, R. J. *Chem. – Eur. J.* **2013,** *19,* 16923–16927. doi:10.1002/chem.201303463
17. Mack, E. T.; Snyder, P. W.; Perez-Castillejos, R.; Bilgiçer, B.; Moustakes, D. T.; Butte, M. J.; Whitesides, G. M. *J. Am. Chem. Soc.* **2012,** *134,* 333–345. doi:10.1021/ja2073033

18. Shan, M.; Bujotzek, A.; Abendroth, F.; Wellner, A.; Gust, R.; Seitz, O.;
Weber, M.; Haag, R. *ChemBioChem* **2011,** *12,* 2587–2598.
doi:10.1002/cbic.201100529

19. Scheibe, C.; Bujotzek, A.; Dernedde, J.; Weber, M.; Seitz, O.
Chem. Sci. **2011,** *2,* 770–775. doi:10.1039/c0sc00565g

20. Krishnamurthy, V. M.; Estroff, L. A.; Whitesides, G. M. Multivalency in
Ligand Design. In *Fragment-based Approaches in Drug Design;*
Jahnke, W.; Erlanson, D. A., Eds.; Wiley-VCH: Weinheim, Germany,
2006; pp 11–53. doi:10.1002/3527608761.ch2

21. Samuel, J.; Sinha, S. *Phys. Rev. E* **2002,** *66,* 050801.
doi:10.1103/PhysRevE.66.050801

22. MacKintosh, F. C.; Käs, J.; Janmey, P. A. *Phys. Rev. Lett.* **1995,** *75,*
4425–4428. doi:10.1103/PhysRevLett.75.4425

23. Kienberger, F.; Pastushenko, V. P.; Kada, G.; Gruber, H. J.; Riener, C.;
Schindler, H.; Hinterdorfer, P. *Single Mol.* **2000,** *1,* 123–128.
doi:10.1002/1438-5171(200006)1:2<123::AID-SIMO123>3.0.CO;2-3

24. Gargano, J. M.; Ngo, T.; Kim, J. Y.; Acheson, D. W. K.; Lees, W. J.
J. Am. Chem. Soc. **2001,** *123,* 12909–12910. doi:10.1021/ja016305a

25. Shoichet, B. K. *J. Med. Chem.* **2006,** *49,* 7274. doi:10.1021/jm061103g

26. Hulme, E. C. *Receptor–Ligand Interactions A Practical Approach;*
Oxford University Press: Oxford, United Kingdom, 1992.

27. Zeng, S.; Baillargeat, D.; Ho, H.-P.; Yong, K.-T. *Chem. Soc. Rev.* **2014,**
43, 3426–3452. doi:10.1039/c3cs60479a

28. Hirst, G. K. *J. Exp. Med.* **1942,** *75,* 49–64. doi:10.1084/jem.75.1.49

29. Oesterhelt, F.; Rief, M.; Gaub, H. E. *New J. Phys.* **1999,** *1,* 6.
doi:10.1088/1367-2630/1/1/006

30. Sauter, N. K.; Hanson, J. E.; Glick, G. D.; Brown, J. H.; Crowther, R. L.;
Park, S. J.; Skehel, J. J.; Wiley, D. C. *Biochemistry* **1992,** *31,*
9609–9621. doi:10.1021/bi00155a013