Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/6/4/eaat0919/DC1

Supplementary Materials for

On the design of precision nanomedicines

Xiaohe Tian, Stefano Angioletti-Uberti, Giuseppe Battaglia*

*Corresponding author. Email: g.battaglia@ucl.ac.uk

Published 24 January 2020, *Sci. Adv.* **6**, eaat0919 (2020)
DOI: 10.1126/sciadv.aat0919

This PDF file includes:

Supporting Information

Binding probability for a monovalent ligand on a multivalent target. We derive an equation for the binding probability of a monovalent ligand to a surface coated with multiple receptors. There are various way to do this, each with a slightly different interpretation. For clarity and to understand the implicit assumptions, we will provide two different derivations, using first a grand-canonical and then a canonical description.

One possible derivation can be done by treating ligand binding as a Langmuir adsorption problem (10) where a surface with N_T adsorbing sites, each of which carries N_R ligands. In this case, we interpret the fraction of sites with a ligand adsorbed, $\theta = N/N_T$, as the adsorption probability. Note that since in the Langmuir adsorption problem each site is independent from each other, this is the same as calculating the probability that a single site has a bound ligand attached. Thus, following the classical Langmuir model, we write

$$\theta = \frac{aq}{1+aq} \tag{S1}$$

where a is the ligand activity which for dilute solutions is given as $a \approx \rho v_0$. With ρ being there the number concentration of ligands $\rho = [L]v_0$, where $[L]$ is the molar concentration of ligands and N_A Avogadro's number and v_0 the so-called binding volume, i.e. the volume in which a ligand must be confined to be considered bound to the site. Note that q is formally the ratio of the partition function for a bound ligand with respect to an unbound, *one once the ligand has been confined within a volume* v_0 around the receptor. In other words it is the change in free-energy if we turned on the interaction potential between the ligand and the receptor, keeping the ligand within the binding volume. If there are N_R receptor per site, and a ligand can only bind one of them at any one time, we have that

$$q = N_R exp(-\beta \Delta G_0) \tag{S2}$$

where $\beta = 1/k_B T$ and ΔG_0 is the ligand-receptor binding free-energy. In order to understand the exact meaning of these quantities, it is important to give a proper definition of this energy. The value of ΔG_0 is linked to the internal partition function of ligands and receptors in the bound state, z_{LR} ,compared to the free ones, z_L and z_R, respectively. More precisely, if M and m are the degrees of freedom of a ligand and receptor, respectively, z_{LR} is the $m + M - 6$ dimensional integral of $exp(-\beta H(\{x\}))$, H being the Hamiltonian of the system, over all the degrees of freedom $\{x\}$ of the bound ligand-receptor complex once its centre of mass has been fixed and the ligand (or, by symmetry, the receptor) is furthermore confined within a volume v_0 around the receptor. Similarly, z_L and z_R are the $M - 3$ and $m - 3$dimensional integral over the degrees of freedom of the ligand and receptor, respectively, once their centre of mass has been fixed in space. With these definitions, we have that

$$exp(-\beta \Delta G_0) = \frac{z_{LR}}{z_L z_R} \tag{S3}$$

How do the microscopic quantities in Eq. S3 relate to experimentally measurable parameters? An equation connecting microscopic and macroscopic quantities can be derived by comparing the microscopic definition of the chemical potential for a bound ligand receptor pair and an unbound ligand or receptor to their thermodynamic definition. The derivation is a bit long and we will not reproduce it here but it can be taken from [Leunissen *et al, J. Am. Chem. Soc. 2010, 132, 1903–1913*]. This provides the following equation

$$exp(-\beta \Delta G_0) v_0 = K_A = \frac{exp(-\beta \Delta G_{bind})}{\rho^\circ} \quad \text{(S4)}$$

where $K_A = K_D^{-1}$ is the experimentally measurable equilibrium association constant for free ligands and receptors in solution (K_D being the dissociation constant), while $\rho^\circ = 1M$ is the standard concentration of ligands and receptors at which the equilibrium constant is measured. The presence of this concentration in the equation connecting the binding energy and the association constant is very important but often creates confusion. We will come back to this aspect later, let us first comment on the meaning of equation S4 and provide a final expression for the binding probability θ. Equation S4 makes explicit that the thermodynamically-defined binding free-energy for the reaction binding reaction between ligand and receptors $L + R \rightarrow LR$, ΔG_{bind}, is not exactly the bond energy ΔG_0 as defined here, and it actually depends on the definition of the chosen value for v_0. This should not really come as a surprise. In an experiment, the equilibrium concentration of bound and unbound complexes can be measured to calculate the equilibrium association constant from the definition $K_A = \frac{[LR]}{[L][R]}$, where $[LR]$ and $[L]$ ($[R]$)are the equilibrium molar concentration of bound ligand-receptor pairs and unbound ligands(receptors), respectively. However, clearly what we define as bound or unbound will depend on what cutoff distance we decide to choose to decide that two reactants A and B are forming an AB complex, which defines our volume v_0. Now by substituting Eqs S3 and S4 into Eq. S1 we finally obtain

$$p_B = \theta = \frac{\rho v_0 N_R exp(-\beta \Delta G_0)}{1 + \rho v_0 N_R exp(-\beta \Delta G_0)} = \frac{\rho N_R K_A}{1 + \rho N_R K_A} \quad \text{(S5)}$$

which is our final expression for Eq. 1 in the main text.

Let us now quickly see where the factor ρ° comes from, which is the source of confusion and can make expression appear dimensionally not correct, if the exact definition or interpretation of concentrations or densities is not given. Binding between a ligand and receptor in solution can be described as a bimolecular reaction $L + R \rightarrow LR$, for which the textbook definition of the equilibrium constant is

$$K_A^p = \frac{[LR]}{[L][R]} = exp(-\beta \Delta G_{bind}) \quad \text{(S6)}$$

This expression makes sense only if K_A^p (the superscript p will be clear in a second) is non-dimensional, which also means that all the concentrations $[L]$, $[R]$ and $[LR]$ must be non-dimensional, which is true if they are intended as *scaled by some reference value*. However, the association constant is typically reported in inverse molar M^{-1}, creating confusion. The difference is what has been sometime referred to as the ''physics'' vs the ''chemistry'' definition of the association constant. When writing Eq. S6, $[L]$, $[R]$ and $[LR]$ in fact are not the molar concentration, but the molar concentration *scaled by the reference value of the ''reactants'' L and R at which the equilibrium association constant is measured*. By convention, this value is $\rho^\circ = 1M$.

If we insist to interpret $[L]$, $[R]$ and $[LR]$ instead as unscaled concentrations, the correct expression for Eq. S6 is

$$K_A^p = \frac{[LR]/\rho^\circ}{[L]/\rho^\circ [R]/\rho^\circ} = \frac{[LR]\rho^\circ}{[L][R]} = exp(-\beta \Delta G_{bind}) \quad \text{(S7)}$$

or in other words

$$\frac{[LR]}{[L][R]} \equiv K_A^c = K_A^p/\rho^\circ = \frac{exp(-\beta \Delta G_{bind})}{\rho^\circ} \quad \text{(S8)}$$

where K_A^c is now what we called the ''chemistry'' definition of the association constant and has indeed the recognisable dimensions of an inverse concentration M^{-1}.

We now shown another possible route to derive Eq. 1, this time using a ''canonical'' approach rather than the grand-canonical one implicit in the derivation of the Langmuir isotherm, Eq S1. We do that because it helps clarify some of the assumptions we are making in defining this binding probability. Consider N monovalent ligands in a volume V and a target bearing N_R receptors. We assume that, if one of the N_R receptors is occupied, then no other ligand can bind to the other $N_R - 1$ receptors present, e.g., due to steric repulsion. We consider dilute solutions, where we can assume ligands in the bulk are almost never in proximity hence their partition function will be that of an ideal gas. Ligands instead gain an energy ΔG_0 upon binding with a receptor (i.e. when they are within a volume v_0 containing the receptor). Given these assumptions, the bound partition function q_B can be calculated considering all configurations where one of the ligands is bound to any of the receptors, and in this case is confined in a volume v_0 around it, whereas the other $N - 1$ ligands are free to float in the rest of the volume $V - N_R v_0$. This is given by

$$q_B = \frac{1}{(N-1)!}(V - N_R v_0)^{N-1}[N_R v_0 exp(-\beta \Delta G_0)] \tag{S9}$$

where the factor $\frac{1}{(N-1)!}$ comes from the fact that we consider indistinguishable ligands, although we can still tell apart whether a ligand is bound or unbound (for example, checking that it is not in the volume v_0 where it can interact with the receptor). Note that if we want to assume that all ligands are distinguishable nothing changes, as long as we do that consistently for both the bound and unbound state, and we would have exactly the same final expression, as shown later.
The unbound partition function q_U is when no ligand is bound, which means they are all free to float outside the volume where they interact with the receptor, thus giving

$$q_U = \frac{1}{N!}(V - N_R v_0)^N \tag{S10}$$

Now the probability to observe the bound state is given by

$$p_B = \frac{q_B}{q_B + q_U} = \frac{\frac{1}{(N-1)!}(V - N_R v_0)^{N-1}(N_R v_0 exp(-\beta \Delta G)}{\frac{1}{N!}(V - N N_R v_0)^N + \frac{1}{(N-1)!}(V - N_R v_0)^{N-1}(N_R v_0 exp(-\beta \Delta G)} =$$

$$= \frac{N\frac{1}{V - N_R v_0} N_R v_0 exp(-\beta \Delta G)}{1 + N\frac{1}{V - N_R v_0} N_R v_0 exp(-\beta \Delta G)} \approx \frac{\frac{N}{V} N_R v_0 exp(-\beta \Delta G)}{1 + \frac{N}{V} N_R v_0 exp(-\beta \Delta G)} = \frac{\rho N_R K_A}{1 + \rho N_R K_A} \tag{S11}$$

where the approximation we used is that $V \gg N_R v_0$, or in other words the binding volume occupied by the receptors is a very small fraction compared to the bulk volume of the solution. Finally, we also replaced$N/V = \rho$ and $v_0 exp(-\beta \Delta G) = K_A$. Clearly, Eq. S11 is exactly the same equation we had derived in the grand-canonical ensemble starting from Eq 1, which was interpreted as the probability for a binding site in the Langmuir picture to be occupied by a ligand (note the analogy between a single target in the canonical ensemble here and an adsorption site in the Langmuir description, each of which bears N_R receptors).
We note here that, as previously stated, if all ligands were considered completely distinguishable, nothing would have changed. In the definition of the bound partition function, Eq. S9, the factor $\frac{1}{(N-1)!}$ would have been replaced by a factor of N, the number of distinguishable configuration with one bound ligand. At the same time, the unbound partition function would have had no $\frac{1}{N!}$factor at all, since this comes from indistinguishability of particles. Overall, this gives again the same final expression for p_B. One curiosity about this derivation is that it makes clear that we are not counting the contribution to the bound state coming from configurations where more than a single ligand is

bound at the same time to the target (which is also an assumption in the derivation of the classical Langmuir model, Eq. S1). In practice, this means the expression we provide is valid for low bond energies / low ligand concentrations, where this probability is small. In any case, due to this approximation, Eq. S9 actually represents an upper bound to the probability of a single target with multiple receptors to have at least one ligand bound to it.

Single bond derivation We now specialise the case of equations 7 and 8 in order to provide an explicit value for the system under consideration. If a single ligand-receptor pair type ζ is present, we obtain

$$\frac{E_{bond}^{\zeta}}{k_B T} = N_{L,\zeta}[ln(p_{L_\zeta} + \frac{1}{2}(1 - p_{L_\zeta})] + N_{R,\xi}[ln(p_{R_\zeta}) + \frac{1}{2}(1 - p_{R_\zeta})] \quad \text{S.12}$$

where $N_{L,\xi}$ is the number of ligands on a nanoparticle that can bind to the corresponding $N_{R,\xi}$ receptors on the surface of the target and the subscript ζ specifies a possible pair. Note that N_L is not necessarily the total number of ligands on the surface of the nanoparticle, as also pointed out by Martinez-Veracoechea and Frenkel, (11) but only those that can, due to the nanoparticle orientation, bind to the surface. In Eq. S.12, $p_{L(R)_\zeta}$is the probability that a ligand(receptor) is unbound, which for a single type of ligand-receptor pair present in the system is given by the following system of coupled equations

$$\begin{cases} p_L + N_R p_R p_L \chi - 1 = 0 \\ p_R + N_L p_R p_L \chi - 1 = 0 \end{cases} \quad \text{S.13}$$

where we have dropped the subscript ζ for simplicity. In writing Equation S13, we assumed that each ligand can bind to each of the receptors on the target *i.e.* the so-called radial topology of binding(7). Other binding topologies can be similarly considered without changing the qualitative features of the results obtained (11). The only physical solutions of the system in Equation S13 is

$$p_L = \frac{(N_L - N_R)\chi - 1 + \sqrt{4N_L\chi + (1 + (N_R - N_L)\chi)^2}}{2N_L\chi} \quad \text{S14}$$

$$p_R = \frac{(N_R - N_L)\chi - 1 + \sqrt{4N_R\chi + (1 + (N_L - N_R)\chi)^2}}{2N_R\chi} \quad \text{S15}$$

whose substitution in Equation S.12 gives the binding free-energy.
For the multiplexed case, equation S.12 is still valid for each ligand-receptor pair if one assumes that each ligand type can only bind a specific type of receptor. In other words, either no or weak cross-interactions with non-cognate receptors must be present. In this case, one can simply solve the same set of coupled equations for each ligand-receptor pair separately and sum up the resulting free-energy contribution over all possible pairs to provide the total binding energy.

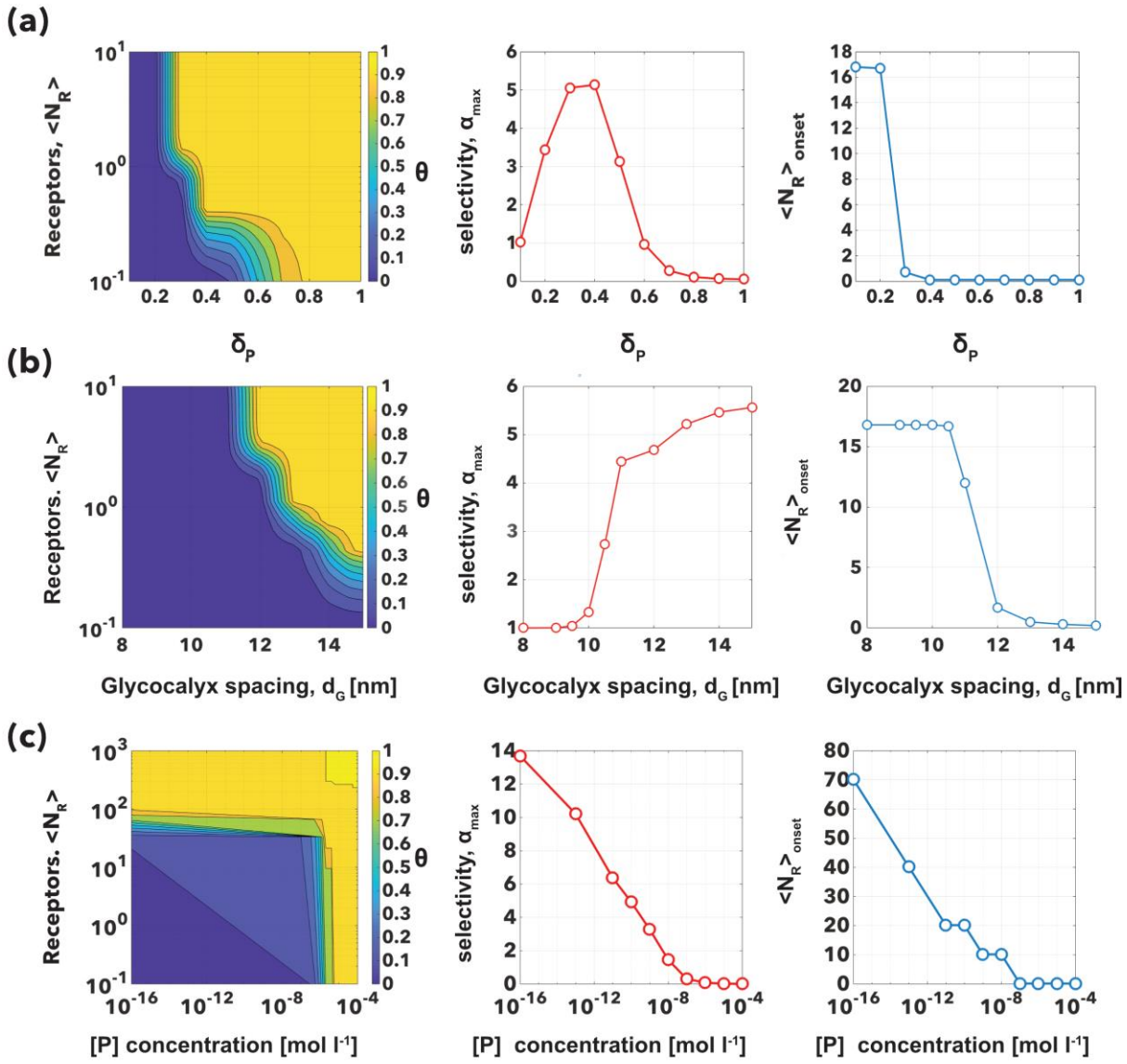

Fig. S1. Scaling principles in superselectivity continued. Heat maps showing the fraction of bound particle θ as a function of the numbers of receptors $\langle N_R \rangle$ and number of the polymer and glycocalyx interference parameter δ_P **(a)** the glycocalyx spacing, d_G **(b)** and particle concentration [P] **(c)**. Each map was analysed to calculate the selectivity α_{max} and the corresponding $\langle N_R \rangle_{onset}$ and the graph of these as function of the varying parameter are reported alongside.

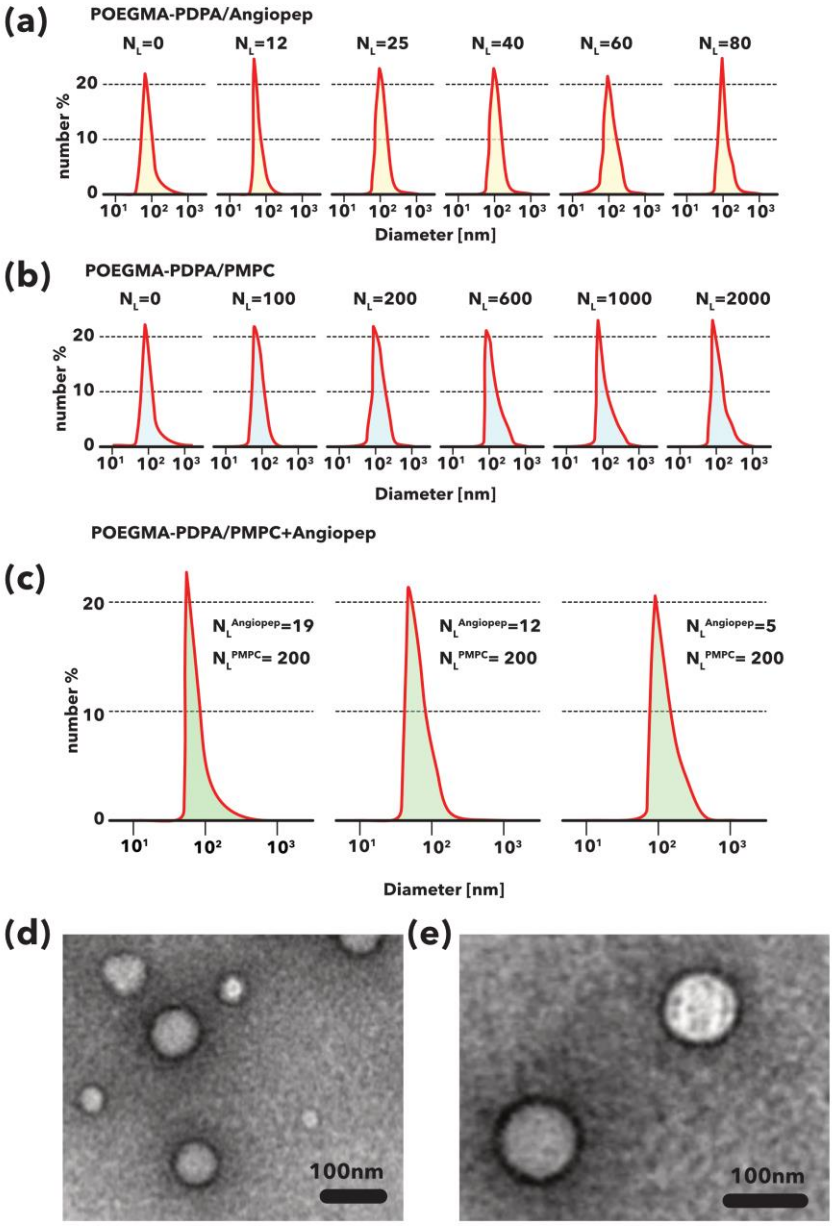

Fig. S2. Polymersome characterization. Particle size distributions measured by dynamic light scattering for POEGMA-PDPA/Angiopep **(a)** POEGMA-PDPA/PMPC **(b)** and POEGMA-PDPA/PMPC + Angiopep **(c)** polymersomes. Representative transmission electron micrographs of POEGMA-PDPA/Angiopep (25 ligands) and POEGMA-PDPA/PMPC (1000 ligands) formulations.

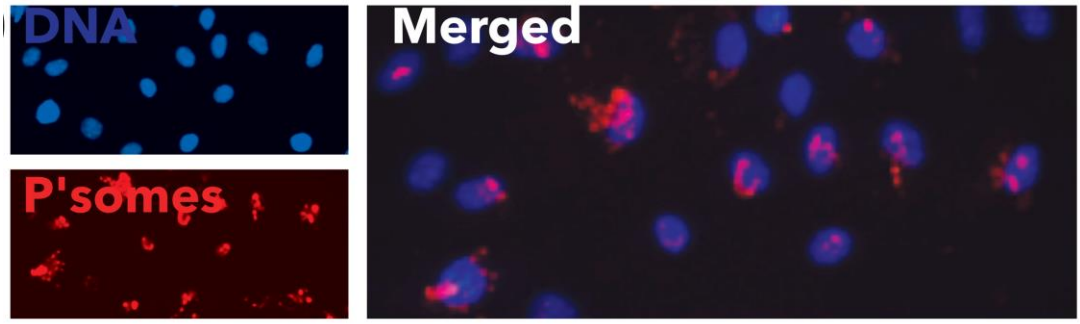

Fig. S3. Polymersome cellular uptake. Micrograph of polymersome bearing 22 angiopep2 ligands uptake in brain endothelial cells (BEnd3) after 1hr incubation. Note the polymersomes were labelled by Cy3 (Red) and the cell DNA by DAPI (Blue)