Design principles for precision targeting

Giuseppe Battaglia\textsuperscript{1,2,}\textsuperscript{*} and Stefano Angioletti-Uberti\textsuperscript{3,4}

\textsuperscript{1}Department of Chemistry, University College London, Christopher Ingold Building, 20 Gordon Street, London WC1H 0AJ, UK
\textsuperscript{2}Department of Chemical Engineering, University College London, Torrington Place, London WC1E 7JE, UK
\textsuperscript{3}Department of Materials, Imperial College London, SW72AZ London, UK
\textsuperscript{4}Beijing Advanced Innovation Centre for Soft Matter Science and Engineering, Beijing University of Chemical Technology, 100099 Beijing, PR China

*Corresponding author: Prof Giuseppe Battaglia, Christopher Ingold Building, University College London, 20 Gordon Street, WC1H 0AJ, London, United Kingdom, Email: g.battaglia@ucl.ac.uk

Tight control on the selectivity of nanoparticles’ interaction with biological systems is paramount for the development of targeted therapies. However, the large number of synthetically tuneable parameters makes it difficult to identify optimal design “sweet spots” without rational guiding principles. Here we address this problem combining super-selectivity theory (SST) with analytical models from soft matter and polymer physics into a unified theoretical framework. Starting from an archetypal system, a polymer-stabilized nanoparticle functionalised with targeting ligands, we use our model to identify the most selective combination of parameters in terms of particle size, brush polymerisation degree and grafting density, as well as tether length, binding affinity and ligands number. We further show how to combine multivalent interactions into multiplexed systems which act holistically as a function of the density of more than one receptor type, so as to achieve binding only when multiple receptors are expressed above a threshold density. We thus propose the design of “bar-coding” targeting approach that can be tailor-made to unique cell populations enabling personalised therapies.

Introduction

Possibly, the most defining feature of a drug is its ability to interact with its biological target as selectively as possible, and indeed most drug discovery tools are fine-tuned to identify those molecules that bind with the highest affinity. This concept goes back to the 19\textsuperscript{th} Century, when Nobel laureate Paul Ehrlich postulated the side-chain theory proposing the existence of receptors and ligands\cite{1}. Selective drugging, popularised as the “magic bullet”, made the fortune of Ehrlich and indeed it is still the cornerstone of modern medicine. Although drug discovery has benefited by many serendipitous discoveries, today it is a highly rigorous process that spans across structural and cell biology, bioinformatics, computational and medicinal chemistry. It is now evolving and merging with -omic technologies to promise personalised therapies\cite{2}. Alongside drug development, we have also advanced our ability to deliver drugs combining the high selectivity of active molecules with molecularly engineered carriers equipped with the necessary attributes to navigate biological environments\cite{3}. A critical element of such a nanomedicinal effort is the introduction of ligands that enable targeting and selectivity to guide carriers across biological barriers. This is now allowing to extend drug discovery to target biological macromolecules that are not accessible via simple passive diffusion such as the inside of cells\cite{4} or the central nervous system\cite{5}.

Today, our ability to create ligands whether for drugging or simply targeting purposes is well advanced and can be extended to almost any biological unit. When the targeted receptor is exogenous to the host, such as the case of infections and poisoning, ligands can be made with selectivity close to ‘Ehrlich’s magic bullet’. However, in most diseases, with cancer being the most exemplary one, the malfunction is associated with receptors that are endogenous and hence expressed by both healthy and diseased cells. Such a promiscuity is the major reason why most drugs come with side-effects and many failed to go through the clinical pipeline. This is also why drug carriers often fail to deliver the drug where it is required and end up adsorbed by the immune system. Yet, such a promiscuous nature is managed with exquisite precision within a living system with molecules, proteins, nucleic acids, and cells interacting with one another with
extremely high selectivity. Such a complex nature of the interaction calls for a complete redesign of our targeting ability to include strategies that indeed allows high selectivity within a highly crowded environment where non-specific interaction can be difficult to overcome.

Historically, the strength of interaction between a given ligand, L, and it receptor, R, is measured by its affinity and this is defined by the same thermodynamic principles that apply to a reversible reaction. The reaction association constant $K_L = k_{on}/k_{off}$, $k_{on}$ and $k_{off}$ being the rates of binding and unbinding, respectively, is defined as the ligand affinity and it has units of inverse concentration. The higher the ligand affinity, the lower the ligand concentration required to saturate its receptor. Affinity can be augmented by combining different ligands into multivalent scaffolds[6] and in such cases, the binding is defined by the term avidity which represents the total effect of the bound units collectively[7]. Multivalent interactions are critical in most biological processes as they allow the translation of weak bonds into strong interaction enabling clustering and signal transduction[8]. Similarly, multivalent interaction is the bread and butter of supramolecular chemistry and often at the core of the design of nanoscale devices[9]. From a theoretical standpoint, the probability of a single ligand to bind to its target can be expressed roughly as $p_{bind} \approx \frac{v_0 n K_L}{(1 + v_0 n K_L)}$, where $n$ is the number of receptors on the target and $v_0$ the so-called binding volume[10]. Thus, probability saturates to 1 for either large number of receptors $n$ or high binding strength. For this reason, high affinity means that a large proportion of ligands will bind to any cells that express the targeted receptors, not just those over-expressing them. This inevitably leads to unwanted interactions, which in the case of chemotherapy, where the final aim is to kill the abnormal cells, can lead to reactions that outweigh the clinical benefits. For multivalent systems the binding affinity has a strong contribution from combinatorial entropy [7]. Noting such a peculiar feature of multivalency, in 2011 Martinez-Veracoechea and Frenkel came up with a very interesting approach, proposing the super-selectivity theory (SST)[10]. This suggests that the combination of low-affinity ligands creates on-off association profiles where the multivalent scaffold saturates the receptors only above a given onset receptor density, whilst does not bind at all at lower densities. Such a scenario is indeed what needed to target cancer cells which often over-express receptors otherwise present in several healthy tissues. SST was proven experimentally in model systems such as supramolecular polymers[11] and multivalent polymers[12, 13]. However, a major limitation to the applicability of SST is that the affinity required to create super-selective profiles is rather low, corresponding to binding energies of order of few $k_B T$’s, where $k_B$ is the Boltzmann’s constant and $T$ is the absolute temperature. Indeed, the receptor and its ligand interact via supramolecular forces such as coulombic forces, hydrogen bonds, aromatic interaction, hydrophilic and hydrophobic effects and van Deer Waals interactions. Although, these are usually weak forces, the range of realistic binding energy is much higher than that required by SST with the lowest limit being the water hydrogen bond of about $8 k_B T$[14] to the strongest biological supramolecular bond known so far, the avidin/biotin complex, with association energy of c.a. $30 k_B T$[15]. In addition to this, as recently demonstrated by Angioletti-Uberti[16], multivalent systems are strongly affected by unspecific interactions induced by the presence of ligands other than the targeted ones and this is exacerbated by the use of low-affinity ligands. In the following, we show how these problems can be solved by combining the general concept of SST theory with principles from soft matter and polymer physics [17] that allow to concurrently modulate the bond-mediated specific interaction and avoid non-specific ones. In doing that, we also show how to achieve multiplexed targeting based on multiple receptor types.

**Theory**

**Rules of engagement in a biological environment.** From blood plasma to interstitial fluid to the cell cytosol, biological liquids are complex aqueous over-saturated solutions comprising several molecules and macromolecules. Typical protein volume fractions can range from 10% in the blood plasma to up to 40% in some cellular organelles. This means that a given protein can be packed with average distance ranging from 6 to 0.3 times its own radius creating dense environments. Such a crowded nature leads to weak non-specific association due to solvation and entropic effects[18]. Also macromolecular crowding means that the same proteins responsible for the crowding have reduced diffusion[18, 19] while metabolites diffusing through the protein network can exhibit enhanced percolation[18]. Finally, the same water that bath the protein can exhibit different properties different from bulk behaviour due to its confinement between the proteins[20]. This creates one of the most, if not the most, complex many-body system. In a first approximation we can propose that any given pair $(ij)$ within a biological system interacts subjected to an inter-particle potential

$$U (r_{ij}) = U^E (r_{ij}) + U^{NS} (r_{ij}),$$

(1)

Assuming a regular hexagonal packing, the ratio between the protein radius, $r_P$ and the inter-protein distance $r_{ij}$ can be derived as a function of the volume fraction $\phi_P$ as

$$\frac{\phi_P \sqrt{T}}{2\pi} = \sqrt{\left(\frac{r_P}{r_{ij}}\right)^2}.$$
where \( r_{ij} \) is the distance between their centre of mass and we assume Boltzmann averaging over all possible mutual orientations. Equation 1 introduces the term \( U^{E} \), which we christen the “Ehrlich potential”, representing a pair-wise, highly selective attraction mediated by ligand-receptor bond formation between the unit \( i \) and its complementary unit \( j \). Any other non-specific interaction can be grouped together into an average potential \( U^{NS} \). Steric hindrance means that \( U^{NS} \) will be always repulsive at \( r_{ij} \sim r_{ij}^{\ast} \), where \( r_{ij}^{\ast} \) is the size of the single unit, and although it can still be attractive at larger distances, colloidal stability requires that its magnitude cannot be much stronger than the thermal energy. Note that whereas generally \( U^{NS} \neq 0 \) for any pair of biological units, an Ehrlich potential only exists when matching ligand-receptor pairs are present. Biological interactions can be tuned by modulating both the specific Ehrlich potential and the non-specific part. Most importantly, any exogenous unit, such as drug or a drug-loaded carrier or even a pathogen, will undergo non-specific interaction with the highly crowded biological environment independently of the magnitude of its Ehrlich potential. In fact, for macromolecular drugs, pathogens and nanoscopic carriers, a larger size determines an enhancement of non-specific binding interactions, which typically lead to fast association with several endogenous proteins. Such a fouling process, also known as opsonisation, often correlates with fast reidance\[21\] and must thus be avoided. On top of this, complex biological organisms posses proteins equipped already to recognise non-self or abnormal self species and hence bestowed with intrinsic Ehrlich potentials that recognise chemical signatures classified by the adaptive immune system as non gratae \[22\]. Hence, the most effective strategy to survive and selectively target a specific biological unit is to associate non-immunogenic Ehrlich pairwise potential, \( U^{E} \) with an non-specific potential \( U^{NS} \) that must be on average repulsive. The latter requirement can be engineered exploiting functional chemical motifs with high affinity with water \[23\]. Such a steric stabilisation cannot be achieved by charged species as their double layer potential is extremely sensitive to other electrolytes present in solution. On the contrary, neutral polar functional groups that form hydrogen bonds can still interact with water orienting it in the close proximity hence creating repulsive steric potentials that can be relatively insensitive to other species, at least within physiological conditions\[24\]. Repulsion, particularly toward proteins, is critical when it comes to biomaterials design\[23\] and indeed these chemical rules in such a contest are known as the Whiteside’s rules\[25\]. Steric stabilisation can be further augmented by polymerising the Whiteside hydrophilic units, adding to the chemical repulsive potentials excluded volume effects associated with the macromolecular nature\[26, 24\]. Such an approach is possibly the most common strategy used in biomaterial design to avoid protein fouling and ensure long-term biocompatibility of surfaces and nanoparticles\[27, 28, 23\]. Typically, the device is coated with an hydrophilic polymer that both satisfies the Whiteside rules and it is not immunogenic. The most used polymer that fulfill such requirements is the poly(ethylene oxide) (PEO) also known as poly(ethylene glycol) (PEG)\[27, 28\]. This is one of the very few synthetic polymers that is generally recognised as safe for most medical applications and it is used routinely in the clinic as adjuvant/coating for several devices and drugs\[29\]. Alternatives to the PEO include poly(vinyl pyrrolidone) (PVP), poly(2-methacryloxyethyl phosphorylcholine) (PMPC), poly(glycerol)s, poly(amino acid)s, polysarcosine, poly(2-oxazoline)s, and poly(N-(2-hydroxypropyl)methacrylamide) \[27, 28\].

**Steric stabilisation.** Now that we have established the ‘rules of engagement’, the quest is to define how to achieve selectivity. To help the discussion and facilitate the model derivation and calculations, while at the same time studying a relevant system, we assume from now on that the exogenous unit is a spherical nanoparticle with core radius \( R \), stabilised by a grafted polymers brush, and equipped with several ligands, see Fig.1a. We also show in Fig.1b the most studied receptors and corresponding natural ligands for cancer targeting whose structure is drawn to scale and the figure showing a large variability in size. We opted to have \( N_L > 1 \) ligands with tether length \( 0 \leq z \leq h \) and, as an example here, we use the transferrin mimic peptide (TMP)\[30\] as ligand to target the human transferrin receptor (HTR). The geometry of such a model system is drawn to scale and shown in Fig. 2a. For ligands’ tethers shorter than the average brush height, binding requires the receptor to be inserted in the brush, thereby feeling the repulsive steric forces produced by the PEO chains. In fact, the receptor feels the equivalent of an osmotic pressure due to the protein volume \( V_P \) being depleted by PEO monomers. This fact is crucial since tuning the tether length gives us a handle to modulate the effective bond energy between ligands and receptors. In this way, we can decrease the effective affinity to values where super-selective behaviour to receptors density can be achieved, avoiding at the same time to increase non-selective interactions. We now provide a firm quantitative basis to these qualitative concepts. To do this, we adapt the Halperin\[31\] model to derive the brush steric potential exerted on a protein (receptor) inserted in the brush as a function of the protein insertion distance from the surface, \( 0 \leq z \leq h \) (see Fig. 2) as:

\[
\beta U_{\text{steric}} = \frac{V_P}{\sigma^{1/2}} \left( 1 - \frac{z}{h} \right)^{2/3}
\]

where \( \beta = (k_B T)^{-1} \). We now assume that: (i) intermolecular forces are larger then intramolecular ones, (ii)
the PEO is in good solvent conditions, and (iii) all the chain ends are localised at the brush edge and thus face the water side. Under these assumptions, we can apply the generalised Zhulina model [32, 33] and define the area per chain $\sigma$ as a function of the insertion distance, $z$ as:

$$\sigma(z) = \sigma_0 \left(1 + \frac{z}{R}\right)^{\gamma - 1}$$  \hspace{1cm} (3)

Note that in our model, as shown in Fig. 2.b, $z$ corresponds to the ligand tether length. Furthermore, in Eq.3, the grafting density at the surface (i.e. $z = 0$) $\sigma_0 = \pi d^2$, with $d$ being the inter-chain distance, and $\gamma$ is a geometrical constant that represents the packing of the chains on the surface. For $\frac{h}{R} \leq (\sqrt{3} - 1)$, $\gamma = \left(\frac{3}{\sqrt{3}} + 1\right)^2$, while for $\frac{h}{R} \geq (\sqrt{3} - 1)$, $\gamma = 3$. For this latter case, the brush scaling does not change with the curvature and it is only controlled by the solvent/polymer interaction which in good solvent conditions means $h \sim N^{1/2}$. According to Zhulina [32], the brush height can be written as:

$$h = R \left[\left(1 + \frac{(\gamma + 2)N}{3R} \left(\frac{\nu a^2}{3\sigma_0}\right)^{1/3}\right)^{3/2} - 1\right]$$  \hspace{1cm} (4)

where $\nu \sim a^3$ is the monomer excluded volume parameter. We define the ratio $\delta = z/h$ as the insertion parameter, and combine equations 2, 3, and 4 to write:

$$\beta U_{\text{steric}} = V_p \left[\sigma_0 \left(1 + \delta \left[\left(1 + \frac{(\gamma + 2)N}{3R} \left(\frac{\nu a^2}{3\sigma_0}\right)^{1/3}\right)^{3/2} - 1\right]\right)^{\gamma - 1} - \frac{\gamma}{(1 - \delta)^{3/2}}\right]$$  \hspace{1cm} (5)

In Fig. 2c, we use Eq. 5 to plot the steric potential as a function of the particle core radius $R$ and $\delta$ fixing the brush polymerisation degree $N = 100$, whereas in Fig. 2d we plot the same quantity as a function of $R$ and $N$ fixing the insertion parameter $\delta = 0.3$. In both cases we assume a relative high grafting density with inter-chain distance $d = 1.05$nm, while we use the HTR structure as shown in Fig. 2a to calculate the protein volume, $V_p = 165.8$nm$^3$ using Chimera and the structural data from the protein data (PDB file : 2NSU). The resulting curve is a considerable energetic barrier where potentials up to 25 $k_B T$ are required to overcome completely the brush repulsion at $\delta = 0$. Most importantly, as expected the potential is insensitive to the curvature for $R \gg h$, whereas for $R \sim h$ only very short tethers creates considerable energetic barriers. As for the polymerisation degree, this has only a minor influence on repulsion in the range $1 < N < 100$, whereas for longer polymers, the probability of steric hindrance decreases. Notwithstanding, the longer the chains are, the larger the particle needs to be in order to create an interfering steric potential. Finally, we know that PEO repels almost any proteins and hence we can comfortably assume that such a potential is very much universal to any receptor.

### Multivalent interactions

Now that we have shown how the polymer brush can be used to tune the binding strength of a single ligand-receptor pair, we need to describe how the possibility to form multiple bonds at the same time affects the overall binding energy of the nanoparticle. In other words, we need a general model to describe multivalent effects. The latter arise from the fact that nanoparticles can use their ligands to bind the cell surface forming many distinct bond arrangements. Each of these constitute a possible micro-state of the system that should be taken into account when calculating the free-energy due to bond formation[7]. Importantly, as first shown by Kitov and Bundle, there is a degeneracy $\Omega$ associated to each micro-state that can strongly contribute to its weight in determining the overall binding free-energy. This is simply due to the fact that this degeneracy translates in an associated entropy, typically named “avidity entropy”, $S_{\text{avidity}} = k \log \Omega$. In calculating the binding energy of a nanoparticle to a receptor-bearing surface, each micro-state should be properly taken into account, including its entropic contribution. Angioletti-Uberti et al showed[34] that when this is done a general analytical formula arises for the free-energy due to bond formation:

$$F_{\text{bond}} = \sum_i \ln(p_i) + \frac{1}{2}(1 - p_i)$$  \hspace{1cm} (6)

where $p_i$ is the probability that a ligand or receptor $i$ is unbound, and the sum is all possible ligand and receptors. The values of $p_i$ are given by the solution of the following set of self-consistent equations:

$$p_i + \sum_{j \in \text{neigh}} p_j e^{-\beta \Delta G_{ij}} = 1,$$  \hspace{1cm} (7)
Figure 1: Schematics of multivalent binding. (a) A spherical nanoparticle decorated with PEO chains where l chains are conjugated to transferrin mimic hexapeptide (TMHP) to target the human transferrin receptor (HTR) (PDB: 2NSU). (b) Example of natural ligands and their respective receptors including cancer associated Epidermal growth factor (EGF) (PDB: 2KV4) and its Epidermal growth factor receptor 1 (EGFR) (PDB: 1IVO) and the Epidermal growth factor receptor 2 (HER2) (PDB: 3U7U). This latter binds also to the synaptic plasticity regulator neuregulin-1 (PDB: 1HAE). Human transferrin (HT) (PDB file: 4X1B) and its receptors HTR. The Tumour necrosis factor alpha (TNF-α) (PDB: 1TNF) and its receptor 1 (TNFR1) (PDB file 1EXT). Folic acid and its receptor the folate receptor-α (PDB file: 4LRH). Immunoglobin G (IgG) (PDB: 1IGT) and one of its natural receptors the the FC-γ-RIIIB receptor a (PDB file: 1FNL). All receptors and ligands are drawn to scale and each structure was reconstructed using the corresponding data from the protein data base (PDB).
Figure 2: Schematics of multivalent binding. (a) The volume that the nanoparticle is allowed to explore while bound to a surface with the schematics of the binding where the ligand/receptor interaction are balanced to the steric potential that the PEO chains exert to the receptor. The steric potential is estimated using equation 2 and plotted as a function of the particle core radius $R$ and the insertion parameter $\delta = t/h$ with the degree of polymerisation $N = 100$ (b) and as function of $R$ and $N$ (c) with $\delta = 0.3$. Note that the inter-chain distance has been fixed $d = 1.05\text{nm}$ and the as protein size I used the volume reported for the human transferrin receptor in the Protein Database (PDB file: 2NSU) $V_p = 165.8\text{nm}^3$. 
one for each ligand or receptor in the system, all coupled together. The sum in the l.h.s of equation 7 runs over all possible neighbours of a ligand binder (i.e. either a ligand or receptor) and \( \chi_{ij} = e^{-\beta \Delta G_{ij}} \) is the bond strength for that specific ligand-receptor pair. In the case where all ligands can bind to a single type of receptor only, and vice-versa, and considering that receptors are mobile on the cell surface, one can take an average over all receptors’ positions and substitute \( \chi_{ij} \) with its averaged value \( < \chi > \), which would then depend only on the type of ligand and receptor considered [35]. In this case, the equations leading to \( F_{\text{bond}} \) can be solved analytically (see the Supporting Information). At this point, it is important to discuss what are the various contributions to the bond energy, since this is crucial to understand how to engineer our targeting system. As shown in [36, 37], the bond energy can be written as:

\[
\Delta G_{ij} = \Delta G^0_{ij} + \Delta G^\text{conf}_{ij},
\]

where \( \beta \Delta G^0_{ij} = -\ln(K^\text{eq}_{ij})^\rho \) is the binding energy from association of ligand \( i \) and receptor \( j \) in solution, as measured by the equilibrium association constant typically reported in experiments \( K^\text{eq}_{ij}, \rho \equiv 1 \) M being the standard concentration. Instead, \( \beta \Delta G^\text{conf}_{ij} \) is a configurational contribution due the constraints imposed by binding [36]. In our system, there are two major contributions we need to include in \( \beta \Delta G^\text{conf}_{ij} \). The first arises due to the mobility of the receptors [35]:

\[
\beta \Delta G^\text{conf,mobile}_{ij} = -\ln \left( \frac{A^\text{bind}_{ij}}{A^\text{free}_{ij}} \right).
\]

This contribution accounts for the fact that in order to interact and bind to a ligand, receptors need to be in its proximity. This limits their position within an area \( A^\text{bind}_{ij} < A^\text{free}_{ij} \), where \( A^\text{free}_{ij} \) is the area they can span in the free, unbound state. Due to its logarithmic dependency, we just need a rough estimates of these areas to calculate \( \Delta G^\text{conf,mobile}_{ij} \). We thus take \( A^\text{bind}_{ij} = 2\pi (\delta h)^2 \), i.e. approximately the area spanned by a rigid ligand, whereas for \( A^\text{free}_{ij} \) we take the surface area exposed by a typical cell, of about \( 400 \mu \text{m}^2 \). The second important contribution to \( \Delta G^\text{conf}_{ij} \) comes from the fact that, as discussed earlier, receptors need to penetrate into the PEO brush and compress it in order to bind a ligand. We thus calculate this latter term assuming that the equilibrium adsorption distance between nanoparticles surface and receptors is the average ligand length, i.e. we set:

\[
\beta \Delta G^\text{conf,PEO}_{ij} = U^\text{PEO} (x = h(1 - \delta)),
\]

Given the free-energy for adsorption defined by Eqs.,6,7, 8,9 and 10, as in Ref.[10] we use a simple Langmuir model to describe the binding of the nanoparticles to a cell, considered as a multivalent surface. By using this model, we implicitly assume that: (i) we have \( M \) surface sites, and the number of receptors per site is randomly distributed around an average value \( r \) following a Poisson distribution ; (ii) nanoparticles adsorb on different sites and hence do not compete for the same receptors and (iii) a surface can only be occupied by one nanoparticle at a time. With these assumptions, we can write the fraction of bound nanoparticles, \( \theta \) as:

\[
\theta = \left\langle \frac{aq}{1 + aq} \right\rangle_p,
\]

where \( \left\langle \right\rangle_p \) is an average over all possible distributions of receptors on the adsorption sites, weighted by their Poisson probability (see [10, 16]). In Eq.11, the term \( a \) is the guest nanoparticle activity and \( q \) is the partition function that describes the interaction between a single nanoparticle and a single host surface. The activity is defined as \( a = v_B e^{\mu^\rho} \), where \( \mu \) is the chemical potential, and \( v_B \) is the volume that each nanoparticle is allowed to explore while bound to a surface. Such a volume can be derived for a spherical particle with radius \( R \) and ligand tether length, \( d \), approaching a surface as:

\[
v_B = \frac{\pi}{3} (3(R + d)^3 - 2R^3)
\]

Under dilute conditions, the activity can be approximated as

\[
a \approx [P]N_A v_B = \frac{\pi}{3}[P]N_A (3(R + d)^3 - 2R^3),
\]

where \([P]\) is the molar concentration of particles in the bulk solution and \( N_A \) the Avogadro number.

The single-site bound state partition function is related to the adsorption free-energy by [16]

\[
q = e^{-\beta F_{\text{bond}}} - 1
\]

where \( F_{\text{bond}} \) is the free-energy due to bond formation, properly summed up over all possible bonding combinations given by Eq. 6. The additional \(-1\) takes into account the fact that the nanoparticle is considered
adsorbed only in the case where at least one bond is present [16, 10]. We can thus combine equations 11, 13 and 14 to obtain:

$$\theta = \left\langle \left( \frac{3}{\pi[P]N_A [3(R + z)^3 - 2R^3]} e^{-\delta P_{\text{bond}} - 1} \right)^{-1} \right\rangle_p$$

Equation 15 associates binding to several design parameters and hence allows to identify the most effective combinations. To facilitate the identification of super-selective regimes we use, from now on, the same selectivity function defined by Martinez-Veracoechea and Frenkel [10] as

$$\alpha = \frac{\delta \log \theta(N_R)}{\delta \log \rho}.$$ 

We christen the receptor number where $\alpha$ takes its maximum value as the onset density $N_{\text{onset}}$ (note that this number is actually the average value per site that controls the Poisson distribution) and the corresponding value of $\alpha(N_{\text{onset}}) \equiv \alpha_{\text{max}}$ as the super selectivity parameter. As discussed previously [10, 16], super-selective binding corresponds to quasi-step-like $\theta(N_R)$ functions where the fraction of bound particles rapidly grows from c.a. zero to c.a. 1 as the numbers of receptor goes above the onset density. Across this threshold value, a minimal change in $N_R$ corresponds to changes of $\theta(N_R) \sim N_R^{\alpha_{\text{max}}}$. Therefore, non-selective binding corresponds to $\alpha < 1$ whereas super-selective profiles will have $\alpha > 1$. The graphs showed in Fig. 3 shows $\theta$ as function of the receptor density as well as the corresponding $\alpha_{\text{max}}$, and $N_{\text{onset}}$ as function of different parameters. Each parameter was optimised to achieve high selectivity and tuneable onset density. As shown in Fig. 3.a, $\theta$ is extremely sensitive to the ligands number, $N_{\text{ligands}}$, with the onset density varying with a normal-like trend with a peak at around 5 ligands and decaying to zero as $N_{\text{ligands}} \to \infty$. It should be stressed that the ligands’ number reported here is not equal to the number of ligands on the nanoparticle, but rather the number of ligands that can engage in binding receptors given a fixed orientation of the nanoparticle, which might depend on geometrical constraints. With our choice of parameters, the selectivity as a function of ligand number is always larger than 1 for $N_{\text{ligands}} > 1$ (we do not report $\alpha_{\text{max}}$ for $N_{\text{ligands}} = 1$ because it can be analytically shown that this is always at most 1, independently from any parameter in the system [10]) and it increases with $N_{\text{ligands}}$ up to a maximum of 4 before slowly decaying for larger numbers. The $N_{\text{onset}}$ instead follows a hyperbolic decay allowing for precise tuning only between 2 and 10 ligands.

For the dependence on the particle radius (shown in Fig. 3.b), the maximum selectivity monotonically decreases with larger particles down to 1 for the largest radius considered. However, the onset $N_{\text{onset}}$ receptor concentration follows a bimodal trend, its variation is rather shallow and hence not critical to titrate the super-selectivity to the desired receptor concentration. The bimodal trend in $N_{\text{onset}}$ is due to two competing effects. On the one side, particles with larger radii have denser brushes and thus stronger repulsive contribution to the bond strength, which effectively decreases with increasing particle size. At the same time, however, the activity coefficient $a$ (and hence selectivity as well as $N_{\text{onset}}$) increase with particle size, leading to the opposite effect and giving the observed non-monotonic dependence. However it is important to emphasise that the radius dependence effectively is mixed with the dependence on the number of active ligands as well as details relating to the ligand tether length. If we increase the radius keeping $N_{\text{ligands}}$ fixed, we effectively reduce the overall grafting density and/or we decrease the ligand length, so that overall the same number of ligands can bind to the target at the same time.

For the ligand binding energy $\Delta G_{\text{on}}$ (Fig. 3.c), we observe no interaction above -$k_B T$ where the binding energy is too low to overcome the steric potential. For lower values, the selectivity parameter $\alpha_{\text{max}}$ decays to non-selective values and plateaus to 1. The $N_{\text{onset}}$ varies considerably between the interaction onset (-$k_B T$) and the -$16k_B T$ allowing the targeting of a large range of receptor concentrations. Similar trend can be observed for the insertion parameter $\delta$ (Fig. 3.d) where both super-selectivity and receptor concentration tuning can be achieved using values between 0 and 0.5. In Fig. 3.e, we plot $\theta$ as function of the receptor density for different receptor’s volumes and while we observe that the selectivity increases with the receptor size this change is not very steep. On the contrary, receptor size changes the onset receptor concentration considerably suggesting that clustering can alter considerably the ability to bind selectively to the desired cells. Finally, both selectivity and onset concentration are strongly affected by the degree of polymerisation Fig. 3.f, with higher values possible at intermediate polymer lengths. Bearing in mind that we fix $N \geq 10$, shorter chains correspond to higher steric potential (see Fig. 2c) and hence propose a very effective way to tune selectivity using macromolecular chemistry. In our discussion, it is important to remind that the exact values of $\alpha_{\text{max}}$ and $N_{\text{onset}}$ as well as the sweet-spots in terms of the optimal parameter range to tune these quantities, depend on the choice of the values for all other parameters involved (we have 6 and we fix 5 of them in each graph). However, the trends observed are not qualitatively affected by this choice.

**Multiplexing** As showed in Fig. 3, we can vary different parameters to achieve the selectivity required for the targeted receptor ensuring that binding occurs above a certain therapeutic threshold. However, the very same physics determining the conditions for super-selectivity makes multivalent targeting of extremely high sensitivity and small changes of some of these parameters can lead to very different outcomes. In principle, this could lead to evolutionary responses in cancer that might simply render the system ineffective. For
Figure 3: Scaling rules in super-selectivity. The fraction of bound particle $\theta$ as a function of the receptor density $\rho$, and the corresponding onset density, $\rho_{\text{onset}}$, and selectivity parameter, $\alpha_{\text{max}}$ varying the number of ligands $l$ (a), the ligand affinity $\Delta G_a$ (b), the ligand insertion parameter $\delta$ (c), the internalisation constant $K_i$ (d), the particle radius, $R$ (e), and the receptor volume $V_P$ (f). Unless varied, parameters were fixed at $l=10$, $\Delta G_a=-11.5k_B T$, $\delta=0.2$, $R=50$nm, $K_i=1$, $V_P=165.8$nm$^3$, $N=100$, and $\sigma_0=0.86$nm$^2$ (i.e. inter-chain distance $d=1.05$nm.).
example, a small mutation in a receptor might make its binding energy towards targeting ligands weaker, shifting the required expression threshold at higher values than those experimentally achievable. In order to make multivalent targeting more robust towards such changes, we propose here to target more than one receptor type at the same time using different ligands. In this way, one can make sure that more complex evolutionary adaptation responses must occur before targeting is made ineffective. In doing this, we can make use of the growing amount of bioinformatic data available about cancer-related receptors and their expression in different cancer lines, making a step closer to fulfilling the “big data” revolution expected in the treatment of cancer [38]. We thus propose the design of nanoparticles comprising $\zeta > 1$ types of multiple ligands where each $i$ type is expressed at numbers $N_i > 1$ on the surface with tether length $z_i$. Each ligand is supposed to target a specific receptor type among those expressed on the surface. Considering that ligands are specific for one type of receptor only, hence competition between different ligand for the same receptor thus not occur, Eqs.6,7 show that the corresponding free energy of multiplexed and multivalent binding, $F_{\text{bond, multi}}$, can be expressed as:

$$F_{\text{bond, multi}} = \sum_{\zeta} F_{\text{bond, } \zeta} \tag{16}$$

where $\zeta$ is an index running over all possible ligand-receptor pairs in the system. The values of $F_{\text{bond, } \zeta}$ is given simply by solving equations 6 and 7 considering a sub-system where only that specific ligand/receptor pairs are present (and hence even in this case an analytical solution is available, see Supporting information). In this multiplexed case, one should also account for the different volumes of the various receptors, leading to $\Delta U_{\text{conf, PEO}} = U_{\text{steric, } \zeta}$ where we explicitly indicated that the steric penalty $U_{\text{steric, } \zeta}$ is different for different ligand-receptor pairs. By taking equation 11 and substituting $F_{\text{PEO}}$ with $F_{\text{PEO, multi}}$ and $F_{\text{bond}}$ with $F_{\text{bond, multi}}$, the adsorption is trivially extended to the multiplexed case:

$$\theta = \left( \frac{3}{\pi [P] N_A [3(R + z)^3 - 2R^3]} \exp \left( -\beta \sum_{\zeta} F_{\text{bond, } \zeta} \right) - 1 \right)^{-1} \tag{17}$$

Note that to write Equation 17 we also implicitly assume, in the calculation of $F_{\text{bond, } \zeta}$, that each ligand is attached to tethers of equal length (i.e. $z_1 = z_2 = ... = z_\zeta$). If we had ligands of different length, the situation would be more complicated, as the system will generally prefer to stay at an intermediate distance from the surface not necessarily corresponding to any tether, simply to maximise its binding strength. In this more complex scenario, however, only quantitative rather than qualitative changes are expected. While multiplexing affects the single binding shifting it toward lower receptor densities, the clear advantage comes from the fact we can engineer holistic binding profiles where nanoparticles bind to surfaces only if they express all the targeted receptors at densities above a given threshold as shown in (Fig. 4) for $\zeta = 2$, 3 and 4. This means that nanoparticles can be designed to target specific cell populations which overexpress unique combinations and compositions of receptors. In other words we can ‘bar-code’ targeting to information gathered from -omic screenings on the specific biological target, hence potentially focus interaction to a single cell population. Note that this is different than the approach developed by Curk et al [39] where it is shown how to design nanoparticles so as to target a specific distribution of receptors in terms of its relative composition. In this latter case, any change around the targeted distribution would decrease the binding probability. In our case, we look for the design conditions where binding would occur when multiple receptor types are expressed above a certain threshold, but anything above that number would still lead to binding, making the system more robust towards any biological fluctuations among the different cell populations.

Conclusions

We present here a general discussion on how exogenous material, whether this is drug and nanoparticle or a pathogen, interact with a complex biological system presenting a very simple potential term that account for specific and unspecific interaction. We use this as ‘rules of engagement’ for the design of selective targeting, we thus derive a model adapting the SST theory to a defined multivalent nanoparticle (see Fig. 2) equipped with realistic binding energies introducing a non-specific repulsive potential by inserting the ligand within a PEO polymer brush. Such a strategy was partially validated by Wang and Dormidontova using Monte Carlo simulation where it was shown that the shielding ligands by long chains leads to the an extra loss of entropy at the onset density [40]. Here we build on this and show, using established models for polymer brush steric...
Figure 4: Multiplexed targeting. The fraction of bound particle $\theta$ as a function of the number of receptors available for each type $r$, at different number of ligands types $\zeta = 2, 3, 4$ increasing from left to write. In each case, we can use our theory to define the regions in space where binding can occur (roughly, those in yellow in the colour map). Note that one can implement a scheme where binding occurs when multiple receptors are concurrently expressed above a certain threshold, whereas a single one would require much higher numbers to be effective. In these plots, the insertion parameter $\delta = 0.5$, the particle radius $R = 50 \text{ nm}$, $\sigma_0 = 0.86 \text{nm}^2$ (i.e. inter-chain distance $d = 1.05 \text{nm}$) and the degree of polymerisation $N = 100$. In order to show possible relevant cases, for the different receptors (A,B,C,D) we have used $\beta \Delta G = -11.5, -14.5$ and $V_p = 165.8, 165.8$ for $\zeta = 2$, $\beta \Delta G = -11.5, 14.5, 11.5$ and $V_p = 165.8, 165.8, 100 \text{nm}^3$ for $\zeta = 3$ and finally $\beta \Delta G = -11.5, -14.5, -11.5, -14.5$ and $V_p = 165.8, 165.8, 100.0, 250.0 \text{nm}^3$ for $\zeta = 4$.

repulsion to proteins, that we can tune the interaction so as to create the low affinity necessary for super-selectivity as showed in Figs.3.a-c. We show that particle size, ligands number, polymer brush length can be computed together with ligand affinity and receptor volume to identify the most efficient formulations to achieve super-selectivity (see Fig.3). Finally, we show that the combination of multiple ligands into a multiplexed systems can indeed create purely super-selective targeting where multiple over-expressed receptors would be required for binding, increasing the robustness of the proposed targeting platform. Overall, the model we present here provides not only a very powerful tool to design personalised nanomedicines but also give important insights into how biological systems can achieve such high selectivity. Indeed one can easily extrapolate from the theory herein presented ‘rules-of-thumb’ to how cells, viruses, bacteria, protein and nucleic acid interact between each other hence adding a powerful tool to the existing system biology approaches.

Acknowledgements

G.B thanks the EPSRC (EP/N026322/1) for funding part of his salary via an Established Career Fellowship. S.A-U thanks the Beijing Advanced Innovation Centre for Soft Matter Science (BAIC-SM) for funding. We also thank Andela Saric, Francesco Gervasio and Giorgio Volpe from UCL as well as Daan Frenkel from University of Cambridge for the critical discussion and to revise the manuscript.

Competing financial interests

The authors declare no competing financial interests.

References


1 Supporting Information

We now specialise the case of equations 6 and 7 in order to provide an explicit value for the system under consideration. For the free-energy due to bond formation between a $\zeta$ ligand-receptor pair, we obtain:

$$F_{\text{bond},\zeta} = N_L \zeta \ln(p_{L\zeta} + \frac{1}{2}(1 - p_{L\zeta})) + r_i \ln(p_{R\zeta} + \frac{1}{2}(1 - p_{R\zeta}))$$

(18)

where $N_L$ is the number of ligands on a nanoparticle that can interact with the surface receptors and $N_R$ the corresponding number of receptors, and the subscript $\zeta$ 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 [10], but only those that can, due to the nanoparticle orientation, bind to the surface. In Eq. 18, $p_{L(R)\zeta}$ 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:

\[
\begin{cases}
 p_L + N_R p_R p_L \chi - 1 = 0 \\
 p_R + N_L p_R p_L \chi - 1 = 0,
\end{cases}
\]

where we have dropped the subscript $\zeta$ for simplicity. In writing Equation 19, 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 [10]. The only physical solutions of the system in Equation 19 is:

\[
p_L = \frac{(N_L - N_R) \chi - 1 + \sqrt{4N_L \chi + (1 + (N_R - N_L) \chi)^2}}{2N_L \chi}
\]

(19)

\[
p_R = \frac{(N_R - N_L) \chi - 1 + \sqrt{4N_R \chi + (1 + (N_L - N_R) \chi)^2}}{2N_R \chi}
\]

(20)

whose substitution in Equation 18 gives the binding free-energy. For the multiplexed case, due to the fact that each ligand can only bind a single receptor, one trivially solve the same set of coupled equations for each ligand-receptor pair present in the system, and sums up the resulting free-energy contribution.