Designing Multivalent Probes for Tunable Superselective Targeting

Galina V. Dubacheva[†]*, Tine Curk[§], Rachel Auzély-Velty[‡], Daan Frenkel[§], Ralf P. Richter[†⊥∥▽]*

[†]CIC biomaGUNE, Paseo Miramon 182, 20009 Donostia - San Sebastian, Spain; [§]Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom; [‡]Centre de Recherches sur les Macromolécules Végétales, CNRS, BP 53, 38041 Grenoble Cedex 9, France; [⊥]Université Grenoble Alpes, DCM, BP 53, 38041 Grenoble Cedex 9, France; [∥]CNRS, DCM, BP 53, 38041 Grenoble Cedex 9, France; [▽]Max-Planck-Institute for Intelligent Systems, Heisenbergstrasse 3, 70569 Stuttgart, Germany.

Classification:

Major category: Physical sciences
Minor categories: Chemistry, Biophysics and Computational Biology

Corresponding authors:

Galina V. Dubacheva (gdubacheva@cicbiomagune.es)
Ralf P. Richter (rrichter@cicbiomagune.es)
Address: CIC biomaGUNE, Paseo Miramon 182, 20009 Donostia - San Sebastian, Spain
Phone: +34 943005329

Keywords: tunability, superselectivity, host-guest multivalent interactions, hyaluronan

Abstract

Specific targeting is common in biology and is a key challenge in nano-medicine. It was recently demonstrated that multivalent probes can selectively target surfaces with a defined density of surface binding sites. Here we show, using a combination of experiments and simulations on multivalent polymers, that such 'superselective' binding can be tuned through the design of the multivalent probe, to target a desired density of binding sites. We develop an analytical model that provides simple yet quantitative predictions to tune the polymer's superselective binding properties by its molecular characteristics such as size, valency and affinity. This work opens up a route towards the rational design of multivalent probes with defined superselective targeting properties for practical applications, and provides mechanistic insight into the regulation of multivalent interactions in biology. To illustrate this, we show how the superselective targeting of the extracellular matrix polysaccharide hyaluronan to its main cell surface receptor CD44 is controlled by the affinity of individual CD44-hyaluronan interactions.

Significance

A basic requirement in biomedical research is the ability to specifically target cells and tissues. Targeting typically relies on the specific binding of a 'ligand' on a tailor-made probe to a 'receptor' on the desired cell/tissue. Conventional probes efficiently distinguish a biological entity displaying the receptor from others that do not, but exhibit limited selectivity when the entities to be distinguished display a given receptor at different densities. Multivalent probes that bind several receptors simultaneously potentially can sharply discriminate between different receptor densities. We demonstrate how such 'superselective' binding can be tuned through probe design to target a desired receptor density, and thus lay the foundation for the rational design of a new generation of analytical, diagnostic and therapeutic probes.

Introduction

Multivalent binding to surfaces plays a key role in material and life sciences,[1] yet its mode of action is still poorly understood.[2] A unique feature of multivalent binding is its potential for 'superselectivity'. Superselectivity implies that the number of multivalent probes that are bound per unit surface area increases faster than linearly with the density of binding sites on the surface.[3] Such a strong dependence allows to target surfaces based on their density of binding

sites, i.e. to sharply discriminate between surfaces displaying binding sites above and below a defined threshold concentration. This concept is promising for the design of novel diagnostic and/or therapeutic probes that target biological entities of interest (e.g. cells, tissues) based on their surface properties: a longstanding biomedical challenge.[4-6]

Identifying the factors that determine multivalent interactions is also crucial for understanding naturally occurring cell-surface binding. A case in point is hyaluronan (HA), an extracellular matrix polysaccharide of importance in biological systems,[7] biomaterials[8] and biomedicine:[9] HA binding to the cell surface was found to be selective to the surface density of the main cell-surface receptor CD44.[10] Thus far, no systematic quantitative study has assessed how the proposed molecular parameters (i.e. receptor density, affinity, HA length, number of accessible HA binding sites[11]) regulate the binding of HA to cell surfaces. Understanding which parameters nature 'tunes' to control HA binding in biological processes such as inflammation[12] and tumour development[13] could provide valuable design principles for synthetic multivalent drugs.

Until recently, quantitative experimental studies of multivalent interactions were challenging. For example, investigations of multivalent binding of polymers to surfaces lacked specificity (e.g. due to polymer/polymer interactions[14,15]) and suffered from poor experimental control over the density of ligands/receptors on polymer[10,11] and surface.[14,15] With a well-defined model system based on host-guest interactions, we recently overcame these limitations and provided the first quantitative experimental demonstration of superselective binding of multivalent polymers to surfaces.[16]

To achieve the desired superselective targeting, it is crucial to match the binding behavior of the multivalent probe with the properties of the target surface.[6] Experimental and theoretical studies have illustrated that multivalent binding of polymeric probes depends on different parameters including their length and/or valency,[17,18] affinity,[19] architecture[20] and flexibility.[21] In the present report, we carry out a systematic quantitative study to uncover how the physico-chemical properties of a multivalent probe can be tuned to efficiently target the desired surface.

We combine experiments, analytical modelling and numerical simulations of multivalent polymers to arrive at a coherent picture of the molecular determinants of superselective binding. To this end, we develop a well-defined and tunable experimental model system based on HA as polymeric scaffold and host/guest interactions. Using this experimental platform, we demonstrate

how superselective binding can be modulated by the molecular design of multivalent polymers. We develop a simple analytical model that can predict how superselective binding depends on the polymer's molecular characteristics. We validate our analytical model against numerical simulations of a coarse-grained polymer model. Finally, we employ the developed model to analyze superselectivity of natural multivalent interactions between HA and CD44.

Results

Experimental model with tunable features. We performed a quantitative characterization of the host/guest interactions between self-assembled monolayers (SAMs) functionalized with guests (ferrocene or adamantane) and hyaluronan (HA) modified with the host β-cyclodextrin (β-CD) (Fig. 1A). We designed this particular model system where hosts (receptors) are attached to multivalent polymers, while guests (ligands) are on surfaces, to suppress undesired non-specific polymer/polymer and polymer/surface interactions.[16] We varied the guest surface density systematically, and investigated the effect of several parameters on the superselectivity of multivalent binding: affinity (i.e. binding strength of individual β-CD/guest interactions), polymer valency (i.e. the number of β-CD moieties per polymer chain), polymer linker (i.e. the linker connecting the HA backbone with β-CD moieties) and polymer concentration (in the diluted regime).

To study the effect of polymer valency, we synthesized HA-β-CD derivatives with different degrees of substitution of HA hydroxyl groups by β-CD ($DS_{\beta\text{-CD}}$, determined as the fraction of β-CD-functionalized disaccharides). We used a two-step synthetic procedure based on the esterification of HA hydroxyl groups with pentenoic anhydride followed by the reaction with a β-CD-thiol derivative (**1**, Fig. 1B).[16,22] This thiol-ene coupling method provides mild and efficient functionalization of polysaccharides, and the $DS_{\beta\text{-CD}}$ can be tuned by varying the HA disaccharide/β-CD-thiol molar ratio.[22] Thus, using 0.09 and 0.30 molar equivalents of β-CD-thiol with respect to the repeating disaccharide unit of HA, we obtained HA-β-CD derivatives with $DS_{\beta\text{-CD}}$ = 3% and $DS_{\beta\text{-CD}}$ = 21%. These are abbreviated in the following as HA_L-β-$CD_{0.03}$ and HA_L-β-$CD_{0.21}$, where the index L reflects the presence of the extended pentenoate linker (around 1 nm contour length) between HA and β-CD. To study the effect of the linker, we also synthesized a construct with a simple amide bond between HA and β-CD (HA_{noL}-β-$CD_{0.04}$), using an acid-amine coupling between HA carboxylic groups and β-CD-NH_2 (**2**, Fig. 1B).

To produce surfaces with different densities of sites for the guest grafting, we formed mixed azide-terminated pegylated SAMs on gold surfaces (Fig. 1C). To produce ferrocene-terminated SAMs (SAM-Fc), we subsequently attached ferrocene using an azide/alkyne click reaction.[16,23] The surface density of ferrocene (Γ_{Fc}) was tuned by changing the fraction of the azide-terminated thiol in solution, and quantified electrochemically from the anodic charge associated with the conversion of Fc to Fc^+ (Fig. 2A, inset).[7,19] Using this approach, Γ_{Fc} was varied from 0.5 to 330 pmol/cm^2, corresponding to root-mean-square distances between neighboring ferrocenes (l_{Fc}) ranging from 18 to 0.7 nm.

To study the effect of the affinity, we prepared adamantane-terminated SAMs (SAM-AD) using click grafting of AD-alkyne (**4**, Fig. 1A) to the azide-terminated SAMs. We chose adamantane because its affinity to β-CD is significantly higher than that of ferrocene: in a phosphate buffer at pH 7, K_d^{AD} = 10 μM [24] and K_d^{Fc} = 200 μM.[25] The reaction of the terminal azide group is expected to change the surface hydrophobicity, which can be followed using contact angle goniometry after the surface modification. We therefore performed comparative contact angle measurements on surfaces before and after grafting of adamantane and ferrocene, respectively (Fig. 2A). In parallel, SAM-Fc samples were characterized electrochemically to determine Γ_{Fc}. Fig. 2 shows that the gold surfaces, which exhibited an initial contact angle of 75±2°, became more hydrophilic after the formation of the pegylated SAM. As expected, the contact angles obtained for the mixed SAMs were intermediate between those of the pure monolayers of each single component,[23] and surfaces became more hydrophobic after their click functionalization with hydrophobic guests. Remarkably, the evolution of the contact angles measured after immobilization of adamantane followed exactly the same trend as that obtained for SAM-Fc (Fig. 2A). This indicates that the click reaction occurs with the same efficiency in both cases and that the adamantane surface density (Γ_{AD}) is similar to Γ_{Fc} for a given thiol ratio. The match in the contact angles for identical surface coverages of SAM-Fc and SAM-AD can be explained by structural (i.e. size) and functional (i.e. hydrophobicity) similarities between Fc and AD. Based on the results of the contact angle measurements, we assumed that $\Gamma_{AD} = \Gamma_{Fc} = \Gamma_{guest}$ and determined Γ_{guest} of SAM-AD electrochemically using SAM-Fc samples prepared in parallel.

The nature of the binding of HA-β-CD derivatives to guest-coated surfaces was first characterized by quartz crystal microbalance (QCM-D, Figs. S2-3). HA binding was initially fast and then slowed down progressively (Fig. S2B-C). After 3 hours of incubation, either no further

binding was observed or additional binding was very slow, suggesting that equilibrium had been attained or was approached. Analysis of the QCM-D data revealed that the HA-β-CD films are typically several 10 nm thick and soft (Fig. S2D), indicating that the surface-bound polymers form loops and/or tails dangling into the solution. The thickness was smaller than or comparable to the polymer's radius of gyration ($R_g \approx 45$ nm [16]), as would be expected for the adsorption of polymers to surfaces.[26] Bound HA-β-CD did virtually not desorb during two hours of rinsing in buffer (Figs. S2B-C); in contrast, free β-CD desorbed completely from SAM-Fc and SAM-AD within a few minutes (Fig. S3A-B). With non-specific interactions being absent (Fig. S3C-D), we conclude that the stable binding of HA chains is the result of specific, multivalent host/guest interactions. Yet, because the individual host-guest interactions are reversible, the chains can dynamically rearrange on the surface thus presumably facilitating equilibration.

To characterize the sensitivity of HA-β-CD derivatives to variations in Γ_{guest}, we quantified the surface density of bound polymers ($\Gamma_{HA\text{-}\beta\text{-}CD}$) by spectroscopic ellipsometry (SE) over the full range of guest surface coverages (Fig. S4). Fig. 3A shows plots of $\Gamma_{HA\text{-}\beta\text{-}CD}$ *vs* Γ_{guest} for the different binding scenarios sketched in Fig. 3B, i.e. **(i)** HA_L-β-$CD_{0.03}$ on SAM-Fc (Fig. 3A, purple), **(ii)** HA_{noL}-β-$CD_{0.04}$ on SAM-Fc (green), **(iii)** HA_L-β-$CD_{0.03}$ on SAM-AD (orange) and **(iv)** HA_L-β-$CD_{0.21}$ on SAM-Fc (blue). These particular systems were chosen to study how the selectivity of polymer binding is affected by the polymer linker (pentenoate linker **(i)** *vs* amide bond **(ii)**), affinity (K_d^{Fc} = 200 μM **(i)** *vs* K_d^{AD} = 10 μM **(iii)**) and polymer valency ($DS_{\beta\text{-}CD}$ = 3% **(i)** *vs* $DS_{\beta\text{-}CD}$ = 21% **(iv)**). In addition, we studied the effect of polymer concentration ($c_{HA\text{-}\beta\text{-}CD}$) by comparing system **(iv)** with **(v)** HA_L-β-$CD_{0.21}$ binding to SAM-Fc at 10-times reduced $c_{HA\text{-}\beta\text{-}CD}$ (cyan). The molecular characteristics of HA derivatives are summarized in Fig. 3C.

Fig. 3A shows that all HA-β-CD derivatives discriminate sharply between surfaces with different guest densities. To evaluate the selectivity of binding towards surface coverage, we use the parameter α (see ref. 3) as a measure of the relative rate of change of the number of bound nanoobjects with the relative increase in the surface density of guests: $\alpha \equiv \mathrm{d}\ln\Gamma_{HA-\beta-CD}\,/\mathrm{d}\ln\Gamma_{guest}$. For systems exhibiting superselectivity, α can reach values higher than one, thus causing a faster-than-linear change in the surface density of bound objects: $\Gamma_{HA-\beta-CD} \propto \Gamma_{guest}{}^{\alpha}$. In the log-log plot in Fig. 3A, straight-line segments with different values of α are shown. As the figure shows, there are superselective regions ($\alpha > 1$) in all studied cases, with similar curve shapes being observed for the different systems. A detailed analysis of

steepest slopes (i.e. for $\Gamma_{HA\text{-}\beta\text{-}CD} < 30$ fmol/cm^2; Fig. 3A, inset) reveals comparable maximal α values for systems **(i), (ii)** and **(iv)**, with a weighted average of 3.4 ± 0.2. A significantly higher value, $\alpha_{max} = 4.6 \pm 0.6$, is obtained in the case of decreased polymer concentration (system **(v)**). The value for system **(iii)** ($\alpha_{max} = 4.9 \pm 1.4$) is in the same range as the other systems, although a large standard error precludes detailed comparison. The experiments show that the superselectively range (i.e. the range of Γ_{guest} where α is maximal) strongly depends on the parameters of the multivalent system. Specifically, changing either the polymer linker (system **(ii)**), the β-CD/guest affinity (system **(iii)**), or the polymer valency (system **(iv)**) can shift the superselectivity range by more than one order of magnitude. To our knowledge, this is the first experimental demonstration that the superselectivity range is tunable over a wide range of densities of surface binding sites by the molecular design of multivalent probes.

Analytical model. To gain physical understanding of the observed behavior, we developed an analytical model which is an extension of the theoretical approach developed for reference system **(i)**.[16] The model is described in detail in the Supporting Information. It assumes the surface to be covered by an array of cubic cells, each of volume $a^3 = (4\pi/3)R_g^3$ and containing n_L ligands ($n_L = \Gamma_{guest} N_A a^2$ guests in our experiments). One or more polymers can bind into a given cell, and the n_R receptors (hosts) per polymer can bind independently to the ligands in the cell. The ligand-receptor binding free energy $F = \ln(K_d a^3 N_A)\, k_B T + U_{poly} + \Delta U_{link}$, which determines the probability that a particular ligand-receptor complex is formed once a polymer is in the cell, contains terms taking care of (i) the host-guest affinity and the confinement of the receptor in the cell ($K_d a^3 N_A$), and (ii) entropic penalties related to the reduced conformational space of the polymer and the linker upon formation of a bond (U_{poly}), where we explicitly consider effects originating from a modification of the linker (ΔU_{link}; relative to a reference system). The free-energy penalty for *i* polymers in a given cell due to polymer-polymer and polymer-surface repulsion is approximated by $U_i = Ai^{9/4} + 0.83 k_B T i$, where *A* (a prefactor in a scaling approximation) is expected to be of order 1 $k_B T$. The model also explicitly considers the gains in combinatorial entropy with increasing guest surface density or polymer valency.

To test the model against an extended range of parameters and to check its predictive ability, we used it to analyze different configurations of the host/guest multivalent system

schematized in Fig. 3B. The model contains the two parameters U_{poly} and A that we cannot determine experimentally. We thus first fitted the reference system **(i)** treating U_{poly} and A as the sole fitting parameters. With a now fully determined set of input parameters, the model was then used to predict the behavior of systems **(iii)**, **(iv)** and **(v)**, using values of K_d, n_R and $c_{HA\text{-}\beta\text{-}CD}$ according to the experimental conditions (Fig. 3C). The magnitude of ΔU_{link} upon variation of the polymer linker is *a priori* unknown, and system **(ii)** was thus fitted using ΔU_{link} as the sole fitting parameter.

The results (Figs. 4A-B and S5) demonstrate that the analytical model can quantitatively reproduce the superselective behavior of the reference system **(i)**, with the two adjustable parameters U_{poly} = 4.6 k_BT and A = 0.35 k_BT having the expected magnitude (i.e. on the order of 1 k_BT). Importantly, the model reproduces all trends in the quality of the superselectivity (i.e. variations in the curve shape and in the maximal α) and in the position of the superselectivity range obtained experimentally for systems **(iii)**, **(iv)** and **(v)**. In particular, the shift in the superselectivity range towards lower Γ_{guest} upon increasing affinity (system **(iii)**) is quantitatively reproduced, while some discrepancy remains upon increasing polymer valency (system **(iv)**). The enhanced superselectivity at reduced polymer concentration (system **(v)**) is also well captured by the model, demonstrating a sizable increase of α_{max} from 3.1 to 4.8 (Fig. 4B). System **(ii)** was also reproduced well, yielding ΔU_{link} = -1.9 k_BT. This shift in free energy between systems **(i)** and **(ii)**, as expected on the order of 1 k_BT, is the quantitative manifestation of the effect of the linker on the bond formation and the ensuing shift in the superselectivity range; detailed consideration of the binding free energy F reveals that replacing the extended pentenoate linker by an amide bond is equivalent to a decrease of K_d by a factor of $\exp(\Delta U_{link}/k_BT) = 6.7$.

The match of theoretical and experimental results (except close to the maximal $\Gamma_{HA\text{-}\beta\text{-}CD}$) is remarkable. It demonstrates that the developed model, although simplified, captures essential features of the complex multivalent interactions between polymers and surfaces and thus can be used to predict the superselective binding behavior of real systems. Specifically, the parameters K_d, n_R, ΔU_{link} and U_{poly} affect exclusively the superselectivity range, in a way that is effectively predicted through the scaling parameter $x_S = \Gamma_{guest} n_R K_d^{-1} \rho_0 e^{-(\Delta U_{link} + U_{poly})/k_BT}$ ($\rho_0 = 1$M is the standard concentration). x_S is derived from the analytical model and is expected to provide faithful predictions as long as the fractions of occupied ligands and receptors are low (see

Supporting Information for details). Indeed, all experimental data sets for a given polymer concentration essentially merge into a single master curve when Γ_{guest} is replaced by x_S (Fig. 4C), illustrating that our experimental systems obey this condition. In contrast, the polymer concentration (c) and the polymer size (R_g) are predicted to affect the range as well as the quality of superselectivity. This combined effect cannot be captured by a simple scaling variable yet quantitative predictions can readily be made using the full analytical model: qualitatively, α_{max} increases with decreasing R_g (Fig. S6A) and c (Figs. 4B and S6B).

Numerical simulations. The analytical model is attractive, because one can readily appreciate the effect of various parameters on superselective binding. However, simplifying assumptions about the polymeric nature of the superselective probe had to be made. In particular, the deformation of the polymer upon binding to the surface was considered exclusively through the constant and empirical parameter U_{poly}. How do these simplifications affect the predictions of the analytical model? To test this, we additionally performed grand canonical Monte Carlo simulations with a soft-blob polymer model that explicitly considers the polymeric properties of the multivalent probe. The simulation approach is described in detail in the Supporting Information and only the main results are presented here.

The simulations revealed multivalent binding of polymers and the formation of loops and tails extending from the surface (Fig. S7), in agreement with the experimental results based on QCM-D analysis (Fig. S2). The polymer surface densities at equilibrium obtained through simulation (Fig. 4D) matched the experimental data at high polymer surface densities well. This lends further support to our earlier conclusion that polymer binding in the experiment reached equilibrium. It also confirms that the overestimation of the maximal $\Gamma_{HA\text{-}\beta\text{-}CD}$ values by the analytical model seen in Fig. 4A arises from the underestimation of U_i at high i (Eqs. S4a-b) due to the neglected polymer deformation, as we had previously proposed.[16] We also note that the maximal slope of the simulated curves (Fig. 4D) tends to be somewhat larger than what is predicted by the analytical model (Fig. 4A). Detailed analysis of the simulation data revealed that U_{poly} decreases with the number of bonds formed per polymer (Fig. S8), i.e. this discrepancy arises from the neglected correlations between hosts within a polymer in the analytical model. It may appear surprising that the analytical model reproduces the quality of superselectivity in the experimental data better than the simulations. We hypothesize that this is fortuitous, due to the weakening of the quality of superselectivity by the finite distribution of polymer molecular

weights in the experiment, which was not considered in the analytical model or in the simulations. The simulations also showed that U_{poly} decreases weakly with n_R (Fig. S8), which explains why the analytical model slightly underestimated the shift in superselectivity range experimentally observed when switching from low to high polymer valency (Fig. 4A; equivalent to the correction of U_{poly} for system (**iv**) by -0.7 k_BT in Fig. 4C).

Nonetheless, predictions of the analytical model for the translation of the superselectivity range along the *x* axis are in perfect agreement with simulations. Indeed, the same shift in the superselectivity range is obtained when the simulated binding free energy F_{sim} is varied by 3 k_BT (from 1 to -2 k_BT; Fig. 4D), equivalent to the 20-fold change in K_d of the experiment (from 200 μM (system **(i)**) to 10 μM (system **(iii)** ; Fig. 4A). In addition, the shape of the curves remains virtually unaffected over a wide range of F_{sim}, a behavior that is correctly reproduced by the analytical model. Taken together, the simulations thus confirm that the analytical model indeed provides a physically reasonable approximation of reality over the experimentally studied parameter space and beyond.

Discussion

Through the choice of the properties of our model systems, our findings are relevant for a better understanding of superselectivity in biological systems. We chose β-CD/AD (K_d = 10 μM) and β-CD/Fc (K_d = 200 μM) complexes that cover the affinity range of HA interactions with its main cell surface receptor CD44 (K_d = 10 – 100 μM).[11] Fig. 5 shows data from the literature on HA binding to CD44-coated microbeads.[10] If we analyse the flow cytometry data of English et al. in the way outlined above, we obtain a value of α_{max} around 4. This observation implies that native HA can superselectively target surfaces displaying CD44, in a similar manner as our β-CD-functionalized HA target guest-covered surfaces. Specifically, an increase in CD44 surface density by a factor as small as two can enhance HA binding up to 16-fold. Moreover, the data sets originating from various CD44 constructs with different HA binding properties can be merged into a single master curve by shifting them along the *x* axis (Fig. 5B), demonstrating that the quality of superselectivity is independent of the CD44 construct. Through the scaling parameter x_S, we can estimate that the affinity varies by approximately 7-fold between the least and most strongly binding CD44 constructs. Interestingly, cells expressing these two constructs at comparable levels were found to exhibit pronounced HA binding and virtual absence of HA

binding, respectively.[10] Moreover, cells have been reported to exhibit a supra-linear increase in HA binding with CD44 expression at the cell surface.[27] These observations indicate that HA can superselectively target CD44 even at the cellular level. HA binding to cells is likely to be more complex than to our model surfaces, e.g. because the CD44 distribution on the cell surface can be heterogeneous.[27,28] Here, the insights gained with our well-defined surfaces provide a new paradigm for understanding cell-surface binding of HA on which future studies can build to understand the physiological implications of superselective targeting.

Thus far, we have adopted a "surface centric" point of view, and described how the properties of multivalent polymers can be tuned to superselectively target a desired surface density of binding sites. This perspective is useful, e.g. for the development of probes to target biological surfaces of interest (e.g. cell surfaces or tissues). The scaling parameter x_S, however, confers a deeper meaning to superselective targeting: it implies that if binding is superselective to one of the parameters contained in x_S, then it is superselective with the same quality to any parameter in x_S. From a "polymer centric" perspective, our theoretical analysis thus allows us to design surfaces that are able to selectively target polymers by their valency. Such surfaces could have interesting applications, e.g. for the development of superselective purification devices featuring microbeads or porous matrices with rationally designed surface functionalities.

In summary, we have developed a highly tunable experimental model system that allowed us to relate the superselective binding behavior of multivalent polymers with the physico-chemical parameters of the interacting system. A simple analytical model, developed and validated against experiments and simulations to reproduce most essential features of the real system, can predict how the design of multivalent polymers is to be tuned to achieve targeting of a surface with a desired superselectivity range and quality. The scaling variable x_S provides a simple rationale for tuning the superselectivity range. Our results demonstrate that, due to superselectivity and tunability, multivalent polymers have the potential to serve as versatile probes in biomedical applications, such as the design of efficient polymeric drugs and drug delivery systems. The presented approach, combining experimental and theory, should also be instructive to develop methods for the rational design of the targeting properties of other types of multivalent probes that are widely used in the field of nano-medicine, such as nanoparticles, nanocapsules or liposomes.[29] Moreover, the insights gained from the present study enhance our

understanding of naturally occurring multivalent interactions, such as those between HA and cell surfaces.

Experimental Methods

Synthesis. HA_L-β-$CD_{0.03}$ and HA_L-β-$CD_{0.21}$ (β-CD derivatives of HA with the pentenoate linker, Fig. 1B, **1**) were synthesized using photochemically induced thiol-ene coupling between HA-pentenoate ($DS_{pentenoate}$ = 0.21) and β-CD-thiol. The synthesis of HA_L-β-CD with $DS_{β\text{-}CD}$ = 0.03 (HA_L-β-$CD_{0.03}$) was done as described previously.[16] To obtain HA_L-β-CD with higher DS (HA_L-β-$CD_{0.21}$), 0.30 molar equivalents of β-CD-thiol (with respect to the repeating disaccharide unit of HA) were used, keeping the whole synthetic procedure the same as for HA_L-β-$CD_{0.03}$. HA_{noL}-β-$CD_{0.04}$ (β-CD derivative of HA without linker, Fig. 1B, **2**) was synthesized using an acid-amine coupling between carboxylic groups of HA and β-CD-NH_2. AD-alkyne (alkyne derivative of adamantane, Fig. 1C, **4**) was synthesized using an acid-amine coupling between 1-adamantaneacetic acid and propargylamine. The purified products were characterized by ^{1}H NMR (AD-alkyne and HA derivatives) and MS (AD-alkyne). A detailed description of synthesis, purification and characterization procedures is given in Supporting Information.

Surface modification. To produce surfaces with tunable guest densities, we first formed mixed SAMs consisting of HS-$(CH_2)_{11}$-EG_6-N_3 and HS-$(CH_2)_{11}$-EG_4-OH on gold surfaces. Prior to SAM formation, the gold-coated surfaces were cleaned by UV/ozone treatment (30 min) and rinsing in ethanol, each gold coating being used only once. After overnight formation, the SAMs were functionalized with guests using an azide/alkyne click reaction. The surface density of guests was tuned by varying the fraction of the azide-terminated thiol in solution used to prepare the SAM. The functionalization of gold surfaces with HS-$(CH_2)_{11}$-EG_4-OH and HS-$(CH_2)_{11}$-EG_6-N_3 and subsequent click-grafting of ethynylferrocene (Fig. 1C, **3**) was performed as described elsewhere.[16] For click modification with AD-alkyne (Fig. 1C, **4**), the same procedure was followed.

Electrochemistry. Electrochemical characterization of ferrocene-terminated SAMs was performed with a conventional three-electrode potentiostat system (Model 620E; CH Instruments, Austin, TX, USA). The working electrode (a gold-coated sensor) was placed at the base of the electrochemical cell (adapted from a Q-Sense Open Module, Biolin Scientific) covered with the electrolyte solution (0.1 M $NaClO_4$, typical volume 3 mL) in which the counter

(platinum) and reference (Ag/AgCl/KCl 3 M) electrodes were immersed. The surface density of ferrocene (Fc) molecules was calculated from the anodic peak (Fig. 2A, inset) using Faraday's equation: $\Gamma_{\mathrm{Fc}} = \frac{Q}{F \times z \times A}$, where Q is the electric charge transferred through the electroactive layer, F = 96485 C/mol is the Faraday constant, z is the number of electrons transferred per molecule ($z_{\mathrm{Fc/Fc+}} = 1$) and A is the surface area of the electrode (A = 1.01 cm^2 as previously determined[23]). Based on repeated measurements, we estimate the standard error in the reproducibility of preparing a desired Fc surface density to be 6%. The error bars provided in Fig. 3A correspond to the standard error in the reproducibility combined with the detection limit of the voltammetric setup (0.3 pmol/cm^2).

Contact angle goniometry. Contact angle measurements were performed on a DSA100 Drop Shape Analyzer (KRÜSS GmbH, Hamburg, Germany). Contact angles were calculated from the average and standard deviation of six measurements, with 3 μL drops of ultrapure water being positioned on different places of the modified gold surfaces.

Spectroscopic ellipsometry (SE). SE measurements were performed in flow mode (20 μL/min) in 10 mM HEPES (pH 7.4) with 150 mM NaCl using a Q-Sense Ellipsometry Module (Biolin Scientific) on gold-coated sensors. The flow module was mounted with the Q-Sense E1 system on a spectroscopic rotating compensator ellipsometer (M2000V; Woollam, Lincoln, NE, USA) and ellipsometric data, Δ and ψ, were acquired over a wavelength range of λ = 380 – 1000 nm at 65 degrees angle of incidence. The multivalent binding was monitored during 3 h of polymer injection and 2 h of buffer rinsing (Fig. S4). All experiments were performed with dilute polymer solutions (i.e. at concentration well below the overlap concentration). Bound polymer masses were determined using the software CompleteEASE (Woollam) by fitting the ellipsometric data to a multilayer model as described previously.[16] The error bars shown in Fig. 3A correspond to the detection limit of the SE setup (0.5 ng/cm^2).

Acknowledgements

This work was supported by the Marie Curie Career Integration Grant "CELLMULTIVINT" (PCIG09-GA-2011-293803) to G.V.D., and the European Research Council Starting Grant "JELLY" (306435) to R.P.R. D.F. acknowledges ERC Advanced Grant 227758 and EPSRC Programme Grant EP/I001352/1. T.C. acknowledges support from the Herchel Smith Fund. L. Yate and J. Calvo (CIC biomaGUNE) are acknowledged for providing

metal surface coatings and help with mass spectrometry analysis, respectively. We thank B. M. Mognetti (ULB, Brussels, Belgium) and O. V. Borisov (IPREM, Pau, France) for fruitful discussions.

Author contributions

Host/guest derivatives were synthesized by G.V.D. and R.A.-V. Experiments were designed by G.V.D. and R.P.R., conducted by G.V.D. and analyzed by G.V.D. and R.P.R. Theory was developed by T.C., D.F. and R.P.R. Simulations were designed by T.C. and D.F., conducted by T.C. and analyzed by T.C. and D.F. Research was directed and manuscript was written by G.V.D. and R.P.R.

Supporting information

This article contains supporting information online at www.pnas.org.

Competing interests

The authors declare no conflict of interests.

References

(1) Mammen, M.; Choi, S.-K.; Whitesides, G. M. *Angew. Chem. Int. Ed.* **1998**, *37*, 2754.

(2) Fasting, C.; Schalley, C. A.; Weber, M.; Seitz, O.; Hecht, S.; Koksch, B.; Dernedde, J.; Graf, C.; Knapp, E.-W.; Haag, R. *Angew. Chem. Int. Ed.* **2012**, *51*, 10472.

(3) Martinez-Veracoechea, F. J.; Frenkel, D. *Proc. Natl. Acad. Sci.* **2011**, *108*, 10963.

(4) Carlson, C. B.; Mowery, P.; Owen, R. M.; Dykhuizen, E. C.; Kiessling, L. L. *ACS Chem. Biol.* **2007**, *2*, 119.

(5) Winau, F.; Westphal, O.; Winau, R. *Microb. Infect.* **2004**, *6*, 786.

(6) Schroeder, A.; Heller, D. A.; Winslow, M. M.; Dahlman, J. E.; Pratt, G. W.; Langer, R.; Jacks, T.; Anderson, D. G. *Nat. Rev. Cancer* **2012**, *12*, 39.

(7) Fraser, J. R. E.; Laurent, T. C.; Laurent, U. B. G. *J. Intern. Med.* **1997**, *242*, 27.

(8) Morra, M. *Biomacromolecules* **2005**, *6*, 1205.

(9) Burdick, J. A.; Prestwich, G. D. *Advanced Materials* **2011**, *23*, H41.

(10) English, N. M.; Lesley, J. F.; Hyman, R. *Cancer Res.* **1998**, *58*, 3736.

(11) Wolny, P. M.; Banerji, S.; Gounou, C.; Brisson, A. R.; Day, A. J.; Jackson, D. G.; Richter, R. P. *J. Biol. Chem.* **2010**, *285*, 30170.

(12) Puré, E.; Cuff, C. A. *Trends Mol. Med.* **2001**, *7*, 213.

(13) Toole, B. P. *Clin. Cancer Res.* **2009**, *15*, 7462.

(14) Ravoo, B. J.; Jacquier, J.-C.; Wenz, G. *Angew. Chem. Int. Ed.* **2003**, *42*, 2066.

(15) Crespo-Biel, O.; Péter, M.; Bruinink, C. M.; Ravoo, B. J.; Reinhoudt, D. N.; Huskens, J. *Chem. Eur. J.* **2005**, *11*, 2426.

(16) Dubacheva, G. V.; Curk, T.; Mognetti, B. M.; Auzély-Velty, R.; Frenkel, D.; Richter, R. P. *J. Am. Chem. Soc.* **2014**, *136*, 1722.

(17) Danial, M.; Root, M. J.; Klok, H.-A. *Biomacromolecules* **2012**, *13*, 1438.

(18) Puffer, E. B.; Pontrello, J. K.; Hollenbeck, J. J.; Kink, J. A.; Kiessling, L. L. *ACS Chem. Biol.* **2007**, *2*, 252.

(19) Mortell, K. H.; Weatherman, R. V.; Kiessling, L. L. *J. Am. Chem. Soc.* **1996**, *118*, 2297.

(20) Dias, R. S.; Pais, A. A. C. C. *J. Phys. Chem. B* **2012**, *116*, 9246.

(21) Dias, R. S.; Pais, A. A. C. C.; Linse, P.; Miguel, M. G.; Lindman, B. *J. Phys. Chem. B* **2005**, *109*, 11781.

(22) Mergy, J.; Fournier, A.; Hachet, E.; Auzély-Velty, R. *J. Polym. Sci., Part A: Polym. Chem.* **2012**, *50*, 4019.

(23) Dubacheva, G. V.; Van Der Heyden, A. l.; Dumy, P.; Kaftan, O.; Auzély-Velty, R.; Coche-Guerente, L.; Labbé, P. *Langmuir* **2010**, *26*, 13976.

(24) Lecourt, T.; Sinaÿ, P.; Chassenieux, C.; Rinaudo, M.; Auzély-Velty, R. *Macromolecules* **2004**, *37*, 4635.

(25) Osella, D.; Carretta, A.; Nervi, C.; Ravera, M.; Gobetto, R. *Organometallics* **2000**, *19*, 2791.

(26) Netz, R. R.; Andelman, D. *Physics Reports* **2003**, *380*, 1.

(27) Perschl, A.; Lesley, J.; English, N.; I., T.; R., H. *Eur. J. Immunol.* **1995**, *25*, 495.

(28) Chen, J.; Pei, Y.; Chen, Z.; Cai, J. *Micron* **2010**, *41*, 198.

(29) Varner, C. T.; Rosen, T.; Martin, J. T.; Kane, R. S. *Biomacromolecules* **2014**.

Figure legends

Fig. 1. Tunable host/guest model system to study multivalent interactions between polymers and surfaces. (A) Schematic representation of HA-β-CD binding to the surface functionalized with guests. (B) Chemical structures of HA-β-CD synthesized by thiol-ene coupling between HA-pentenoate and β-CD-thiol (**1**) and by acid-amine coupling between carboxylic groups of HA and β-CD-amine (**2**). (C) Chemical structures of azide-terminated pegylated SAM and alkyne-derivatives of the guests ferrocene (**3**) and adamantane (**4**). (D) Table of tuned parameters.

Fig. 2. Grafting guest molecules to azide-terminated SAMs. (A) Contact angles measured before and after click functionalization of azide-terminated SAMs with Fc (SAM-N_3, black and SAM-Fc, green) and AD (SAM-N_3, red and SAM-AD, blue) plotted versus Γ_{Fc} determined electrochemically. Each data point corresponds to the mean ± standard error calculated from 6 measurements performed on different positions on the same sample. In addition, sample-to-sample reproducibility was tested for a selected SAM, prepared using 20% azide-terminated thiol: standard errors in the contact angle were 3.6% (42.0 ± 1.5 degrees) before and 3.9% (58.6 ± 2.3 degrees) after click functionalization with AD. The dotted black lines are guides for the eye. An example of a cyclic voltammogram recorded to determine Γ_{Fc} is given in the inset (the measured sample is indicated by the arrow). (B) Representative examples of images showing water drops on surfaces at different modification stages (the measured samples are indicated by the arrow in A).

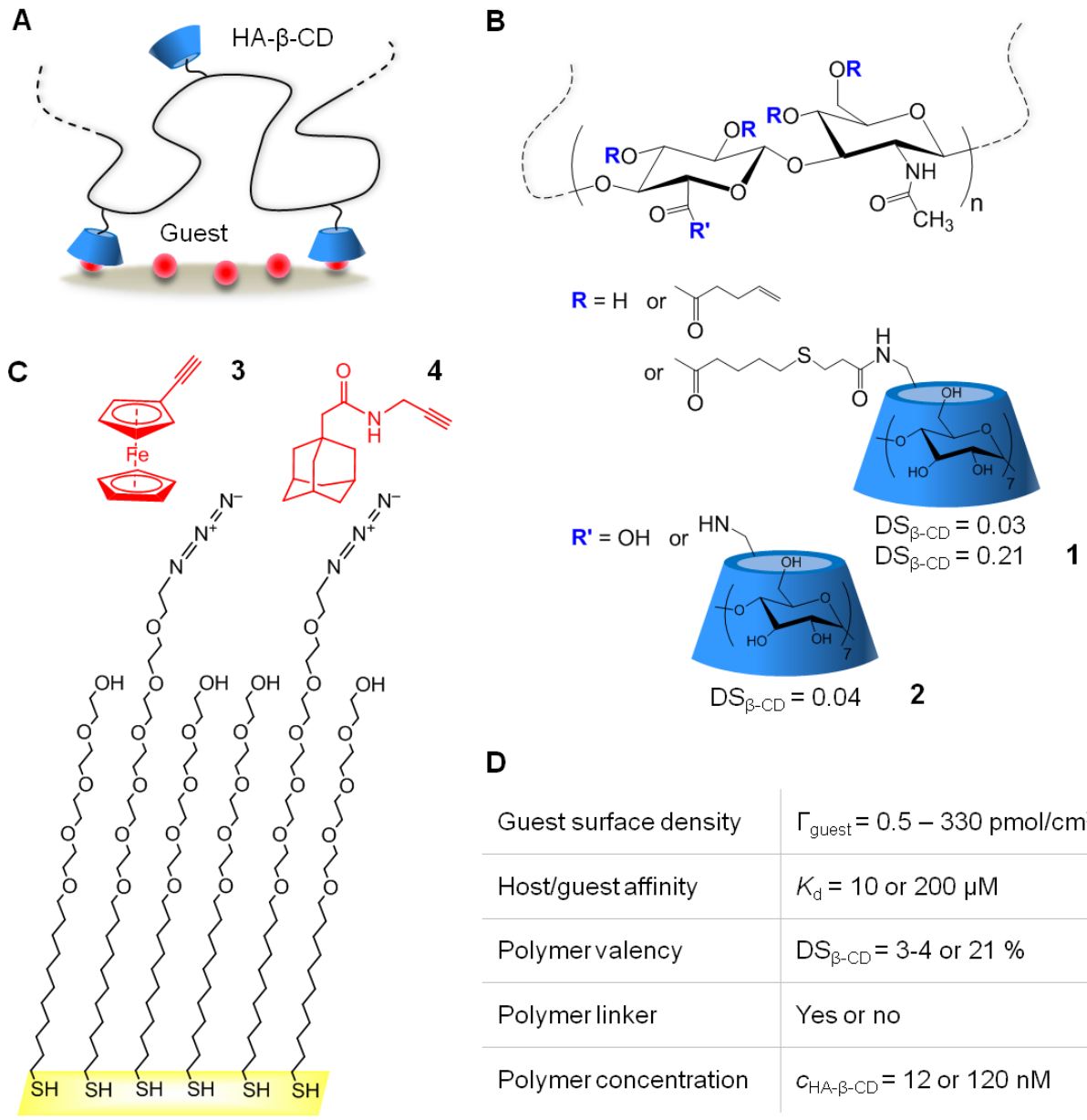
Fig. 3. Experimental characterization of HA-β-CD selectivity to Γ_{guest}. (A) The multivalent binding was monitored by SE during 3 hours of polymer injection and 2 hours of buffer rinsing (Fig. S4). $\Gamma_{HA\text{-}\beta\text{-}CD}$ (determined at the end of the incubation procedure) *vs* Γ_{guest}, is plotted in the form of error bars, with dotted lines connecting data points. $\Gamma_{HA\text{-}\beta\text{-}CD}$ was determined for the binding of HA_L-β-$CD_{0.03}$ to SAM-Fc (purple), HA_{noL}-β-$CD_{0.04}$ to SAM-Fc (green), HA_L-β-$CD_{0.03}$ to SAM-AD (orange) and HA_L-β-$CD_{0.21}$ to SAM-Fc (blue, cyan). The data were obtained at $c_{HA\text{-}\beta\text{-}CD}$ = 120 nM (purple, orange, green, blue) and 12 nM (cyan). For the lowest $\Gamma_{HA\text{-}\beta\text{-}CD}$, no significant binding could be detected and only an upper limit is given corresponding to the sensitivity of our SE setup (0.5 ng/cm^2). The slopes corresponding to α = 1, 3 and 5 are shown to facilitate data interpretation; the inset display maximal α values ± standard errors, estimated for

each system through fitting of the three data points showing lowest HA-β-CD binding to a power law. Conditions: buffer – 10 mM HEPES (pH 7.4), 150 mM NaCl, flow rate – 20 µL/min, T = 23 °C. (B) Schematic representation of the different multivalent binding scenarios. (C) Characteristics of HA derivatives calculated from the weight-averaged molecular weight of HA (M_{HA} = 357 kg/mol) and $DS_{\beta\text{-CD}}$: average polymer molecular weight ($M_w^{\text{HA-}\beta\text{-CD}}$), average polymer contour length between adjacent β-CD moieties ($l_{\beta\text{-CD}}$) and average number of β-CDs per polymer chain ($n_{\beta\text{-CD}}$).

Fig. 4. Theoretical characterization of HA-β-CD selectivity to Γ_{guest}. (A) The solid lines are fits with the analytical model for systems **(i)** and **(ii)** and predictions for systems **(iii)**, **(iv)** and **(v)** (see Fig. 3B); the experimental $\Gamma_{\text{HA-}\beta\text{-CD}}$ *vs* Γ_{guest} data (from Fig. 3A) are shown in the form of error bars for comparison. The fit for reference system **(i)** resulted in $A = 0.35\, k_B T$ and $U_{poly} = 4.6\, k_B T$. These values together with the other known experimental parameters (Fig. 3C) were used to predict the behavior of the systems **(iii)**, **(iv)** and **(v)**. Fitting system **(ii)** with A and U_{poly} kept fixed resulted in $\Delta U_{link} = -1.9\, k_B T$. (B) Dependencies of the selectivity parameter α on Γ_{guest} for the model data in A. (C) Scaling behaviour of polymer binding isotherms. Experimental data presented in Fig. 4A, all at identical polymer concentration, are re-plotted as a function of the scaling variable $x_S = \Gamma_{guest} n_R K_d^{-1} \rho_0 e^{-(\Delta U_{link} + U_{poly})/k_B T}$; for system (**iv**), U_{poly} was additionally reduced by 0.7 $k_B T$ to account for a weak dependence of U_{poly} on n_R that is neglected by the analytical model. The dotted black line is a guide for the eye. (D) Binding isotherms obtained by numerical simulations for low valency polymers for a range of interaction parameters F as indicated. Parameters: n_R = 27, N_b = 20, $c_{\text{HA-}\beta\text{-CD}}$ = 120 nM. Each data point corresponds to a single simulation run, with dotted lines connecting data points.

Fig. 5. HA binds superselectively to its main cell surface receptor CD44, and the superselectivity range is tuned by the affinity of CD44 constructs. (A) Experimental data adapted from Fig. 4A in English et al.[10] The authors coated microspheres with protein constructs containing the extracellular domain of various CD44 constructs with distinct glycosylation levels. Relative surface densities of CD44 and HA were quantified as the median fluorescence intensity (MFI) arising from fluorescein, conjugated to the CD44-specific monoclonal antibody IM7 and sparsely to HA, respectively; MFI of microspheres without fusion protein was subtracted to adjust for background signal. CD44 constructs were taken from cell lines that are

constitutively active (Δ), inducible (□) and inactive (○) with regard to HA binding. In addition, CD44 constructs from inducible (■) and inactive (●) cell lines were incubated with neuraminidase, an enzyme that affects the glycosylation of CD44 thus enhancing HA binding. Inset: schematic representation of HA binding to the surface functionalized with CD44 constructs. (B) Data in A with CD44 surface densities re-scaled by a factor 1/γ, with γ adjusted (as indicated) such that all data merge into a master curve. The master curve displays $\alpha_{max} \approx 4$, demonstrating that the binding of HA to its native cell surface receptor is highly superselective. According to the predictions of the scaling parameter x_S (Eq. S8), γ equals the change in the affinity K_d' of the different CD44 constructs for HA relative to the affinity K_d^{ref} of the reference construct (here chosen to be the inducible CD44 without neuraminidase treatment), i.e. $\gamma = K_d'/K_d^{ref}$. Overall, the K_d varies by almost 7-fold, corresponding to a 7-fold shift in the superselectivity range.



D

Guest surface density	Γ_{guest} = 0.5 – 330 pmol/cm^2
Host/guest affinity	K_d = 10 or 200 μM
Polymer valency	DS$_{β-CD}$ = 3-4 or 21 %
Polymer linker	Yes or no
Polymer concentration	$c_{HA-β-CD}$ = 12 or 120 nM

Fig. 1.

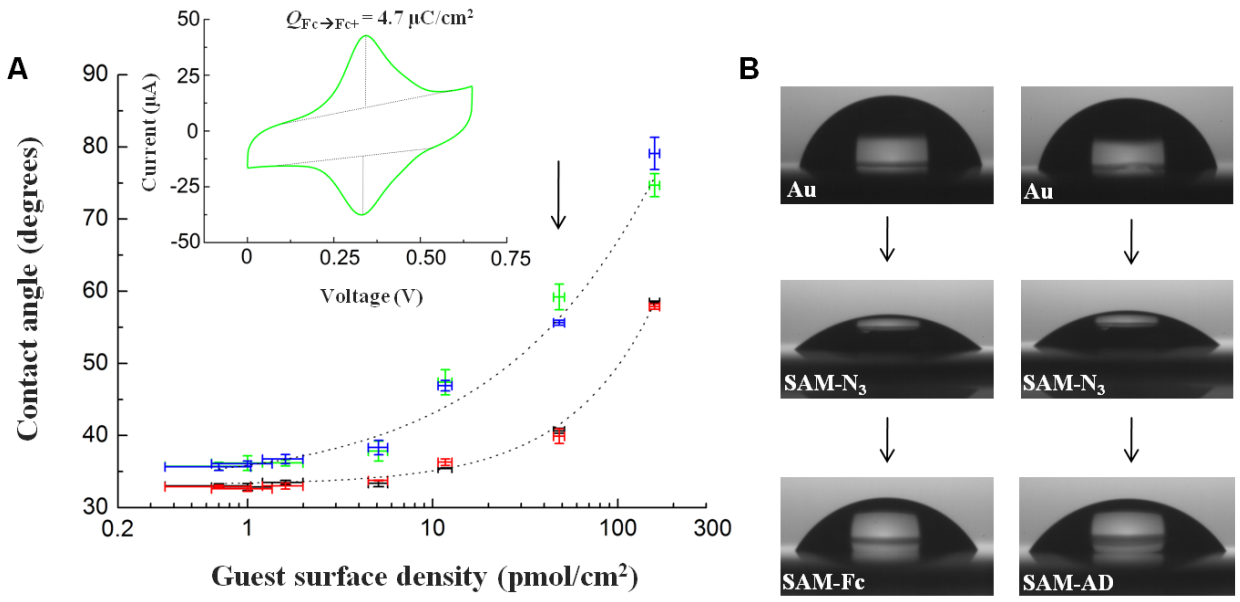
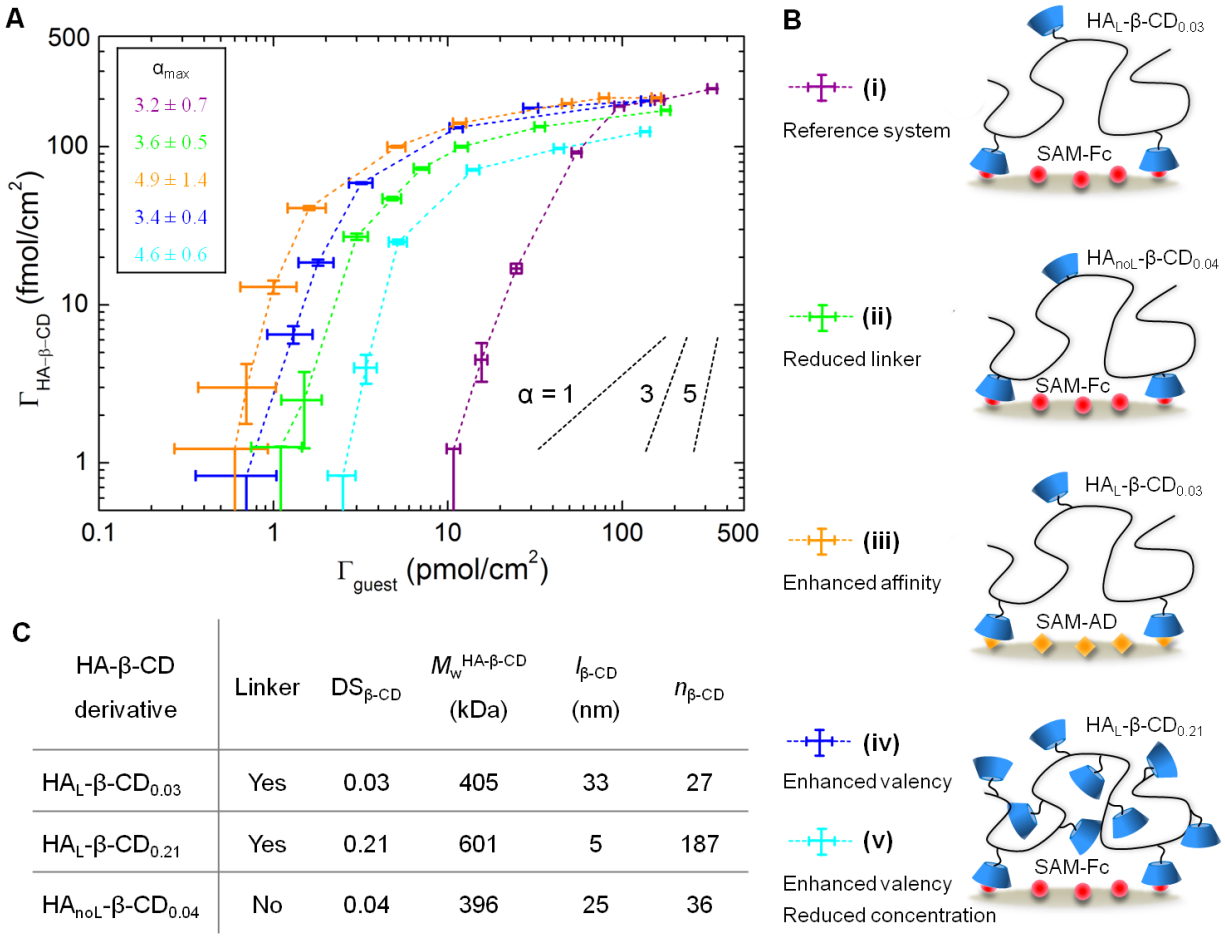

A
B
Contact angle (degrees)
Guest surface density (pmol/cm²)
Current (μA)
Voltage (V)
$Q_{Fc \rightarrow Fc+}$ = 4.7 μC/cm²
Au
Au
SAM-N3
SAM-N3
SAM-Fc
SAM-AD

Fig. 2.



HA-β-CD derivative	Linker	$DS_{\beta\text{-CD}}$	$M_w^{\text{HA-β-CD}}$ (kDa)	$l_{\beta\text{-CD}}$ (nm)	$n_{\beta\text{-CD}}$
HA_L-β-$CD_{0.03}$	Yes	0.03	405	33	27
HA_L-β-$CD_{0.21}$	Yes	0.21	601	5	187
HA_{noL}-β-$CD_{0.04}$	No	0.04	396	25	36

Fig. 3.

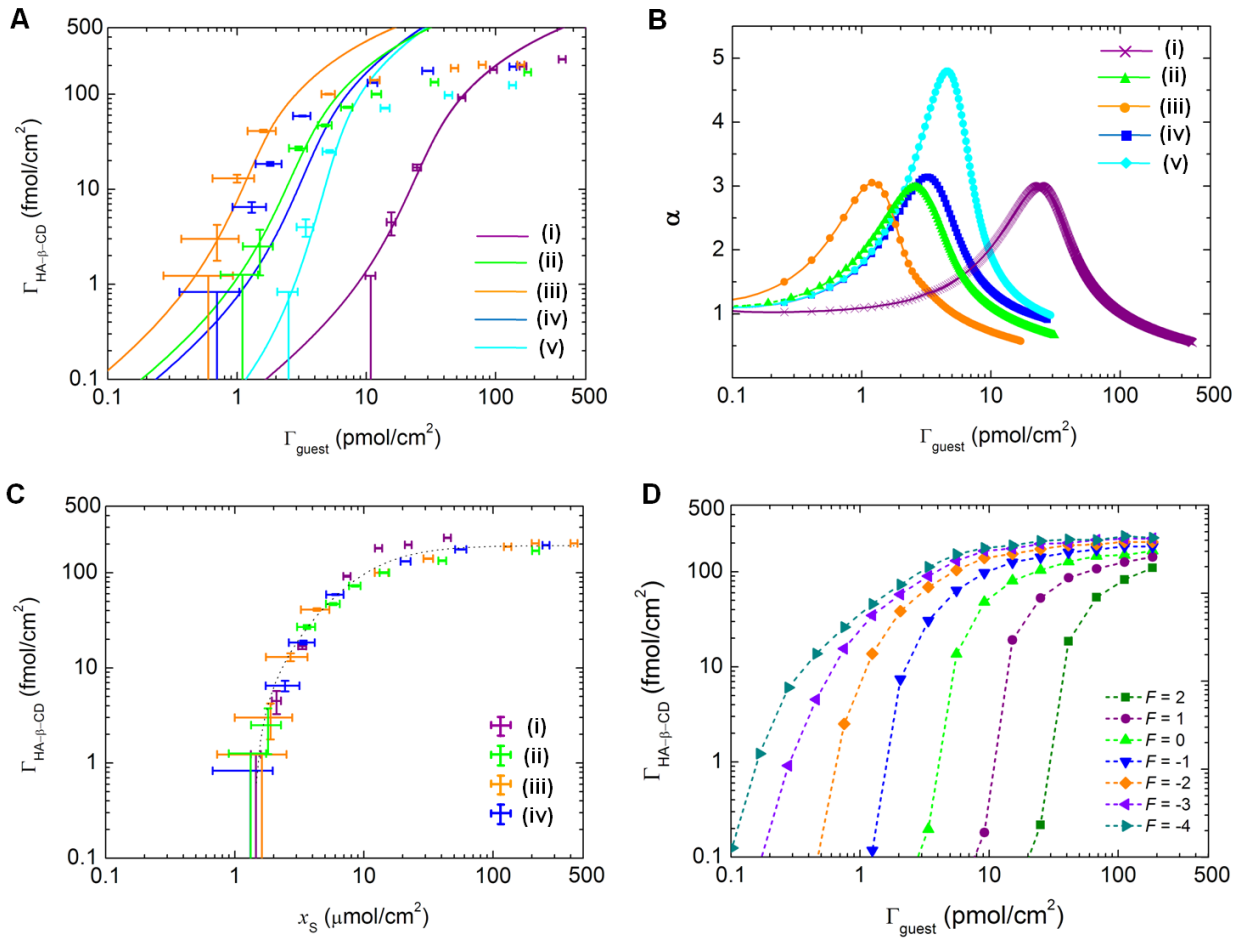


Fig. 4.

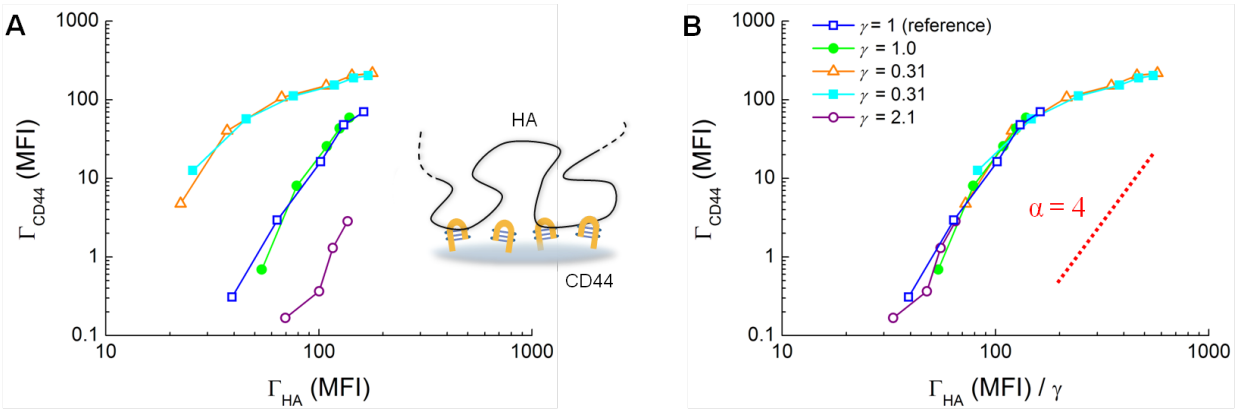

A
1000
100
10
1
0.1
10
100
1000
Γ_{CD44} (MFI)
Γ_{HA} (MFI)
HA
CD44
B
γ = 1 (reference)
γ = 1.0
γ = 0.31
γ = 0.31
γ = 2.1
α = 4
Γ_{HA} (MFI) / γ

Fig. 5.

Designing Multivalent Probes for Tunable Superselective Targeting

Galina V. Dubacheva[†]*, Tine Curk[§], Rachel Auzély-Velty[‡], Daan Frenkel[§], Ralf P. Richter[†⊥∥▽]*

[†]CIC biomaGUNE, Paseo Miramon 182, 20009 Donostia - San Sebastian, Spain; [§]Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom; [‡]Centre de Recherches sur les Macromolécules Végétales, CNRS, BP 53, 38041 Grenoble Cedex 9, France; [⊥]Université Grenoble Alpes, DCM, BP 53, 38041 Grenoble Cedex 9, France; [∥]CNRS, DCM, BP 53, 38041 Grenoble Cedex 9, France; [▽]Max-Planck-Institute for Intelligent Systems, Heisenbergstrasse 3, 70569 Stuttgart, Germany.

Supporting information

Table of content

1. Experimental Methods

1.1. Materials. Hyaluronan (HA) with a weight-averaged molecular weight of 357 kg/mol was purchased from Lifecore Biomedical (Chaska, MN, USA). The overlap concentration c^* for this HA sample in phosphate-buffered saline at 25 °C and pH 7.0 (derived from the intrinsic viscosity [η] assuming that $c^*[\eta] \approx 1$)[1] is approximately 1.5 mg/mL. Ethynylferrocene, 6-monodeoxy-6-monoamino-β-cyclodextrin hydrochloride (β-CD-$NH_3^+Cl^-$), 1-adamantaneacetic acid, propargylamine, 3-mercaptopropionic acid (MPA), *N,N*-diisopropylethylamine (DIEA), hydroxybenzotriazole (HOBt), *N,N′*-diisopropylcarbodiimide (DIC) and 4-pentenoic anhydride were purchased from Sigma-Aldrich. HS-$(CH_2)_{11}$-EG_4-OH and HS-$(CH_2)_{11}$-EG_6-N_3 (EG – ethylene glycol) were purchased from Prochimia (Sopot, Poland). 2-Hydroxy-1-[4-(2-hydroxyethoxy)phenyl]-2-methyl-1-propanone (Irgacure 2959) was kindly provided by Ciba Specialty Chemicals (Basel, Switzerland). Deuterium oxide was obtained from SDS (Vitry, France). The water used in all experiments was purified to achieve a resistivity of 18.2 MΩ cm. Gold-coated QCM-D sensors with a 4.95 MHz resonance frequency (QSX301) were purchased from Biolin Scientific (Västra Frölunda, Sweden).

1.2. Synthesis. The ^{1}H NMR spectra were obtained at 25 °C (AD-alkyne) or 80 °C (HA derivatives) using an ADVANCE III HD spectrometer operating at 400 MHz. Mass spectra were obtained using a LCT Premier™ XE time-of-flight (TOF) mass spectrometer with a ZSpray™ source for electrospray ionization (ESI).

1.2.1. Synthesis of AD-alkyne (alkyne derivative of adamantane, Fig. S1, **1**) was synthesized using an acid-amine coupling between 1-adamantaneacetic acid and propargylamine. To a solution of 1-adamantaneacetic acid (500 mg, 2.57×10^{-3} mol, 1 molar equivalent) in dry DMF, HOBt (695 mg, 5.14×10^{-3} mol, 2 molar equivalents), DIC (1592 μL, 1.028×10^{-2} mol, 4 molar equivalents) and propargylamine (181 μL, 2.83×10^{-3} mol, 1.1 molar equivalents) were successively added. The resulting mixture was stirred under nitrogen atmosphere at room temperature overnight. After evaporation of DMF, the products were dissolved in diethyl ether and purified by silica gel chromatography (1:1 diethyl ether/petroleum ether). AD-alkyne (445 mg, 75% yield) was characterized by ^{1}H NMR and MS.

The chemical shifts δ (in ppm) for ^{1}H NMR ($CDCl_3$) corresponding to the characteristic signals are 4.10-3.95 (dd, 2H, C**H$_2$** next to C≡CH), 2.25-2.17 (t, 1H, C≡C**H**), 1.95-1.90 (broad, 2H, C**H$_2$** next to adamantane), 2.00-1.90 (broad, 3H, C**H** of adamantane), 1.75-1.50 (m, 12H, C**H$_2$** of adamantane). *m/z* found in TOF MS was 232.1702, while $[M+H]^+$ calculated for $C_{15}H_{21}NO$ is 232.17013.

1.2.2. Synthesis of HA_{noL}-β-$CD_{0.04}$ (β-CD derivative of HA without linker, Fig. S1, **2**) was synthesized using an acid-amine coupling between the carboxylic groups of HA and β-CD-$NH_3^+Cl^-$. HA (50 mg, 1.25×10^{-4} mol disaccharides, 1.0 molar equivalent) was dissolved in 3 mL of water, and the resulting mixture was kept at 4 °C under continuous stirring overnight. Then β-CD-$NH_3^+Cl^-$ (73 mg, 6.25×10^{-5} mol, 0.5 molar equivalents), sulfo-NHS (8 mg, 3.75×10^{-5} mol,

0.3 molar equivalents) and EDC (24 mg, 1.25×10^{-4} mol, 1.0 molar equivalent), each dissolved in 1 mL of water, were successively added during stirring (molar equivalents are given with respect to the repeating disaccharide unit of HA). Next, the pH was adjusted to 4.75 and the solution was stirred for 4 h. The product was purified by diafiltration and recovered by freeze-drying. HA_{noL}-β-$CD_{0.04}$ (53 mg, 95% yield) was characterized by ^{1}H NMR.

The chemical shifts δ (in ppm) for ^{1}H NMR (D_2O) corresponding to the characteristic signal intensities are 5.10-4.85 (broad, 7H, C**H** anomeric protons of β-CD), 4.70-4.20 (2 d, 2H, C**H** anomeric protons of HA disaccharide), 3.90-3.10 (m, 10H of HA disaccharide, 8 C**H** and 1 C$\mathbf{H_2}$; 42H of β-CD heptasaccharide, 28 C**H** and 7 C$\mathbf{H_2}$) and 2.10-1.70 (s, 3H of HA disaccharide, C$\mathbf{H_3}$). $DS_{β\text{-}CD}$ = 0.04 was determined by digital integration of the signals at 2.10-1.70 and 5.10-4.85 ppm as described previously.[2]

1.2.3. Synthesis of HA_L-β-$CD_{0.03}$ and **HA_L-β-$CD_{0.21}$** (β-CD derivatives of HA with the pentenoate linker, Fig. S1, **3**) were synthesized using photochemically induced thiol-ene coupling between HA-pentenoate ($DS_{pentenoate}$ = 0.21, Fig. S1, **4**) and β-CD-thiol (Fig. S1, **5**) as described previously.[2] Briefly, β-CD-thiol was synthesized using an acid-amine coupling between β-CD-$NH_3^+Cl^-$ and MPA. HA-pentenoate was prepared by esterification of the HA hydroxyl groups using 4-pentenoic anhydride. HA-pentenoate was first dissolved in water, and the resulting mixture was kept at 4 °C under stirring overnight. Then, Irgacure 2959 (a water soluble photoinitiator, 1 mg/mL) and β-CD-thiol were added during stirring. The concentrations of HA and Irgacure 2959 in the resulting solution were 3 mg/mL and 1 mg/mL, respectively. The reaction mixture was exposed to UV light (λ = 365 nm) at room temperature for 5 min. The product was purified by diafiltration (Amicon YM30) with ultrapure water. The purified product was recovered by freeze-drying and characterized by ^{1}H NMR spectroscopy. The synthesis of HA_L-β-CD with $DS_{β\text{-}CD}$ = 0.03 (HA_L-β-$CD_{0.03}$) was described previously.[2] To obtain HA_L-β-CD with higher DS (HA_L-β-$CD_{0.21}$), 40 mg of HA-pentenoate (9.56×10^{-5} mol disaccharides) was mixed with 35 mg of β-CD-thiol (2.86×10^{-5} mol, 0.30 molar equivalents with respect to the repeating disaccharide unit of HA). The mass of the purified HA_L-β-$CD_{0.21}$ was 58 mg (90% yield). HA_L-β-$CD_{0.21}$ was characterized by ^{1}H NMR.

The chemical shifts δ (in ppm) for ^{1}H NMR (D_2O) corresponding to the characteristic signal intensities are 5.10-4.85 (broad, 7H, C**H** anomeric protons of β-CD), 4.70-4.20 (2 d, 2H, C**H** anomeric protons of HA disaccharide), 3.90-3.10 (m, 10H of HA disaccharide, 8 C**H** and 1 C$\mathbf{H_2}$; 42H of β-CD heptasaccharide, 28 C**H** and 7 C$\mathbf{H_2}$), 2.9-2.2 (m, 8H of pentenoate linker, 4 C$\mathbf{H_2}$), 2.10-1.70 (s, 3H of HA disaccharide, C$\mathbf{H_3}$) and 1.6-1.1 (m, 4H of pentenoate linker, 2 C$\mathbf{H_2}$). $DS_{β\text{-}CD}$ = 0.21 was determined by digital integration of the signals at 2.10-1.70 and 5.10-4.85 ppm as described previously.[2]

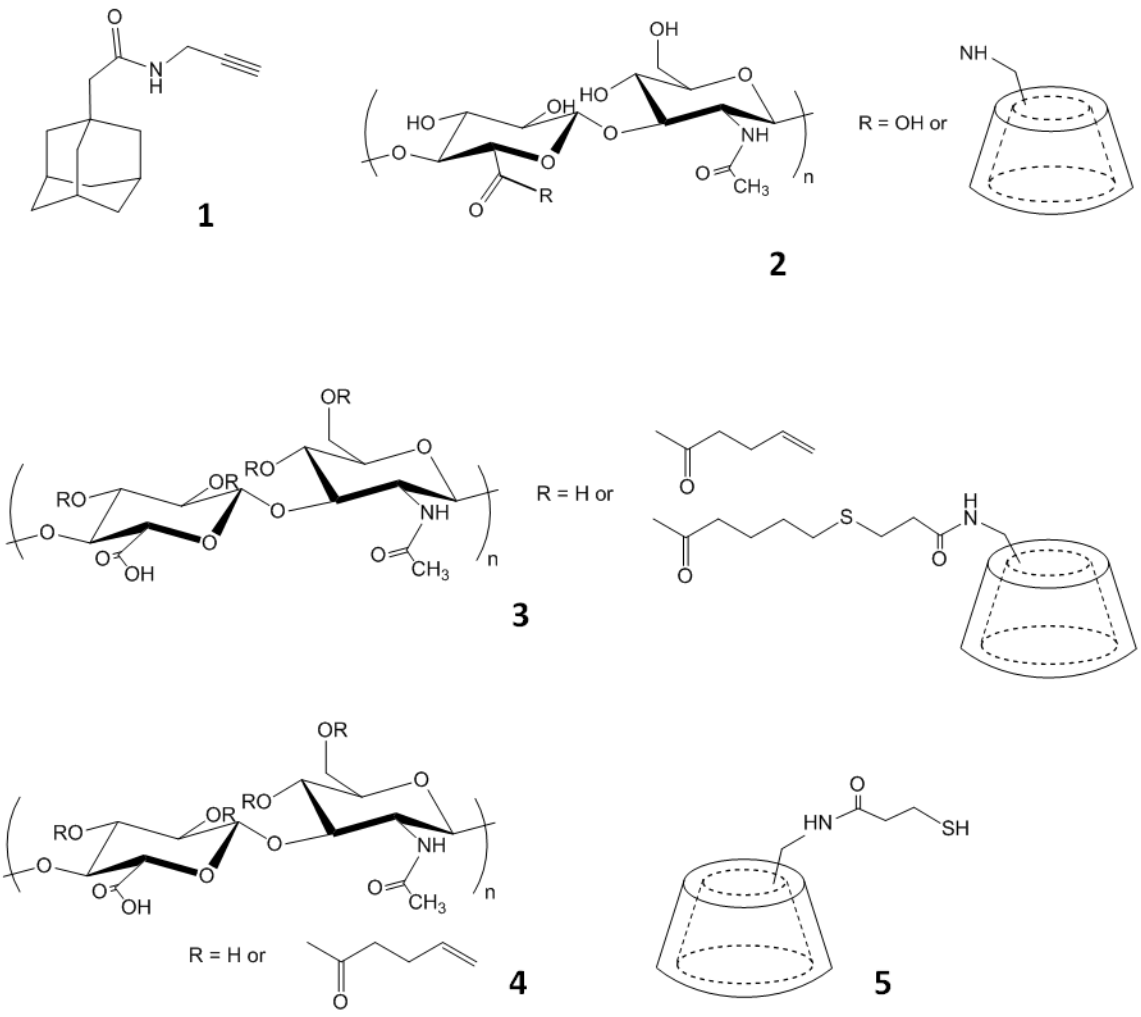
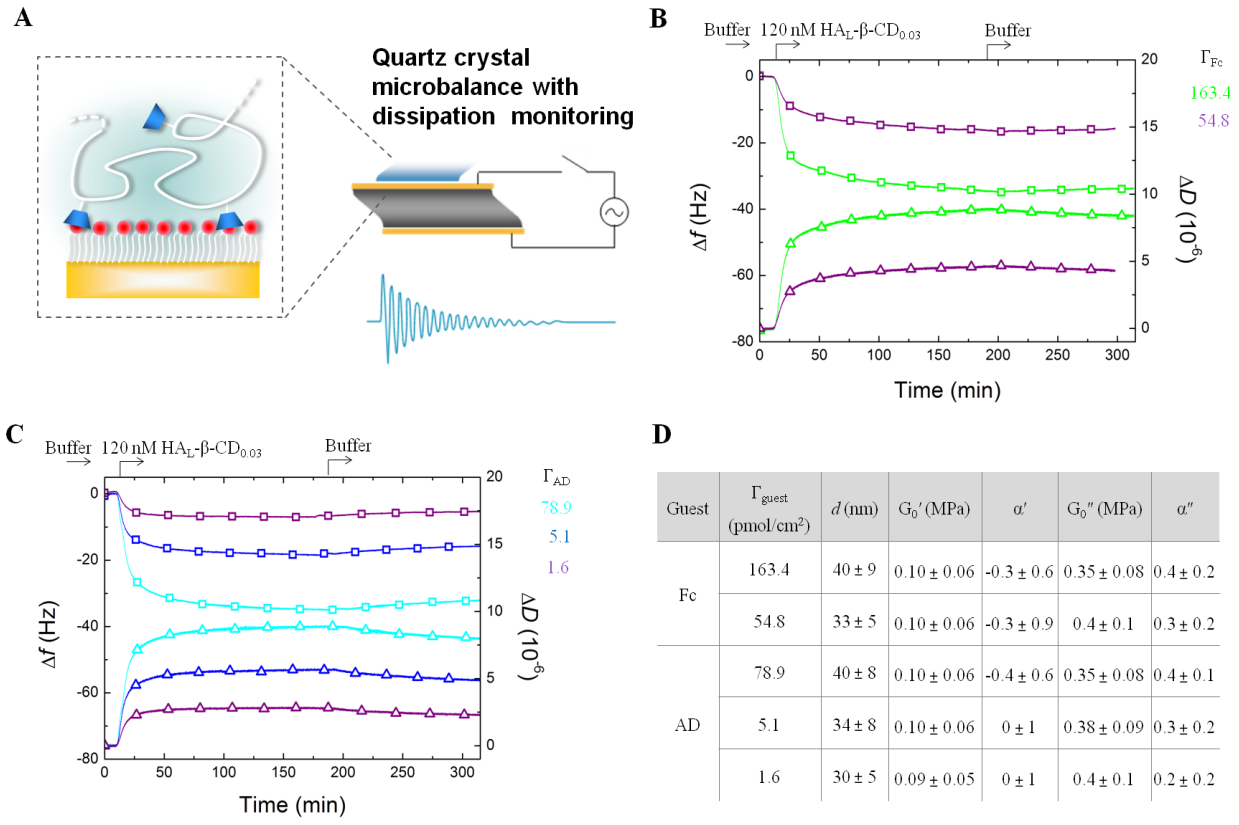


Fig. S1. Chemical structures of AD-alkyne (1), HA_{noL}-β-$CD_{0.04}$ (2) and HA_{L}-β-$CD_{0.03\ /\ 0.21}$ (3). HA_{L}-β-$CD_{0.03\ /\ 0.21}$ derivatives were synthesized through esterification of HA hydroxyl groups using pentenoic anhydride followed by the reaction of HA-pentenoate (4) with a β-CD thiol (5).

1.3. Surface characterization

1.3.1. Characterization of HA binding and HA film morphology by quartz crystal microbalance with dissipation monitoring (QCM-D). QCM-D measurements were performed in flow mode on gold-coated QCM-D sensors, using a Q-Sense E4 system equipped with Flow Modules (Biolin Scientific). Overtones j = 3, 5, 7, 9, 11, and 13 were recorded, corresponding to resonance frequencies of approximately 15, 25, 35, 45, 55 and 65 MHz. Changes in dissipation, ΔD, and normalized frequency, $\Delta f = \Delta f_j/j$, for j = 5 are presented; any other overtone would have shown similar trends. The thickness d and viscoelastic properties of HA-β-CD films were determined by fitting of the QCM-D data to a continuum viscoelastic model,[3] implemented in the software QTM[4] (D. Johannsmann, Technical University of Clausthal, Clausthal-Zellerfeld, Germany). The fitting procedure was described in detail previously.[5] Viscoelastic properties were parameterized in terms of the shear storage modulus $G'(f)$ and the shear loss modulus $G''(f)$.

The frequency dependencies of the storage and loss moduli were assumed to follow power laws within the measured range of 15 to 65 MHz, with exponents α' and α", such that $G(f) = G_0 (f/f_0)^{\alpha}$, respectively, with f_0 set to 15 MHz. The exponents were confined to the ranges $0 \leq \alpha' \leq 2$ and $-1 \leq \alpha'' \leq 1$, i.e. the ranges physically reasonable according to polymer theory.[6] The film density was assumed to be 1.0 g/cm^3. Reduced χ^2-values were verified to be close to 1.0 (thus confirming a good fit), and the indicated errors correspond to a confidence level of one standard deviation.



Guest	Γ_{guest} (pmol/cm^2)	d (nm)	G_0' (MPa)	α'	G_0'' (MPa)	α''
Fc	163.4	40 ± 9	0.10 ± 0.06	-0.3 ± 0.6	0.35 ± 0.08	0.4 ± 0.2
	54.8	33 ± 5	0.10 ± 0.06	-0.3 ± 0.9	0.4 ± 0.1	0.3 ± 0.2
AD	78.9	40 ± 8	0.10 ± 0.06	-0.4 ± 0.6	0.35 ± 0.08	0.4 ± 0.1
	5.1	34 ± 8	0.10 ± 0.06	0 ± 1	0.38 ± 0.09	0.3 ± 0.2
	1.6	30 ± 5	0.09 ± 0.05	0 ± 1	0.4 ± 0.1	0.2 ± 0.2

Fig. S2. HA-β-CD binding to SAMs monitored by QCM-D. (A) Schematic representation of the QCM-D setup. Representative QCM-D responses (Δf - lines with squares, ΔD – lines with triangles) for HA-β-CD film formation are shown for HA$_L$-β-CD$_{0.03}$ binding to SAM-Fc (B) and to SAM-AD (C). Conditions: $c_{HA\text{-}\beta\text{-}CD}$ = 120 nM, buffer – 10 mM HEPES (pH 7.4) with 150 mM NaCl, T = 23 °C, flow rate = 20 µL/min; arrows on top of the graphs indicate the start of injection with a new solution; guest surface densities (in pmol/cm^2) for the respective measurements are indicated by numbers in matching colours on the right of the graphs. Elevated dissipation shifts indicate that the HA films are soft. (D) Table of HA-β-CD film thickness (d) and viscoelastic properties (G_0', α', G_0" and α'') , determined by viscoelastic modelling of data at the end of the HA-β-CD binding process.

To confirm that non-specific interactions do not interfere with host/guest binding, several control measurements were performed. Fast and complete desorption was detected for β-CD monovalently bound to SAM-Fc/AD (Fig. S3A-B). As desired, no shifts were observed upon exposure of HA_p (i.e. HA modified with pentenoate, the precursor for the synthesis of HA-β-CD) to SAM-Fc/AD and HA_L-β-$CD_{0.03}$ to a guest-free SAM (Fig. S3C-D).

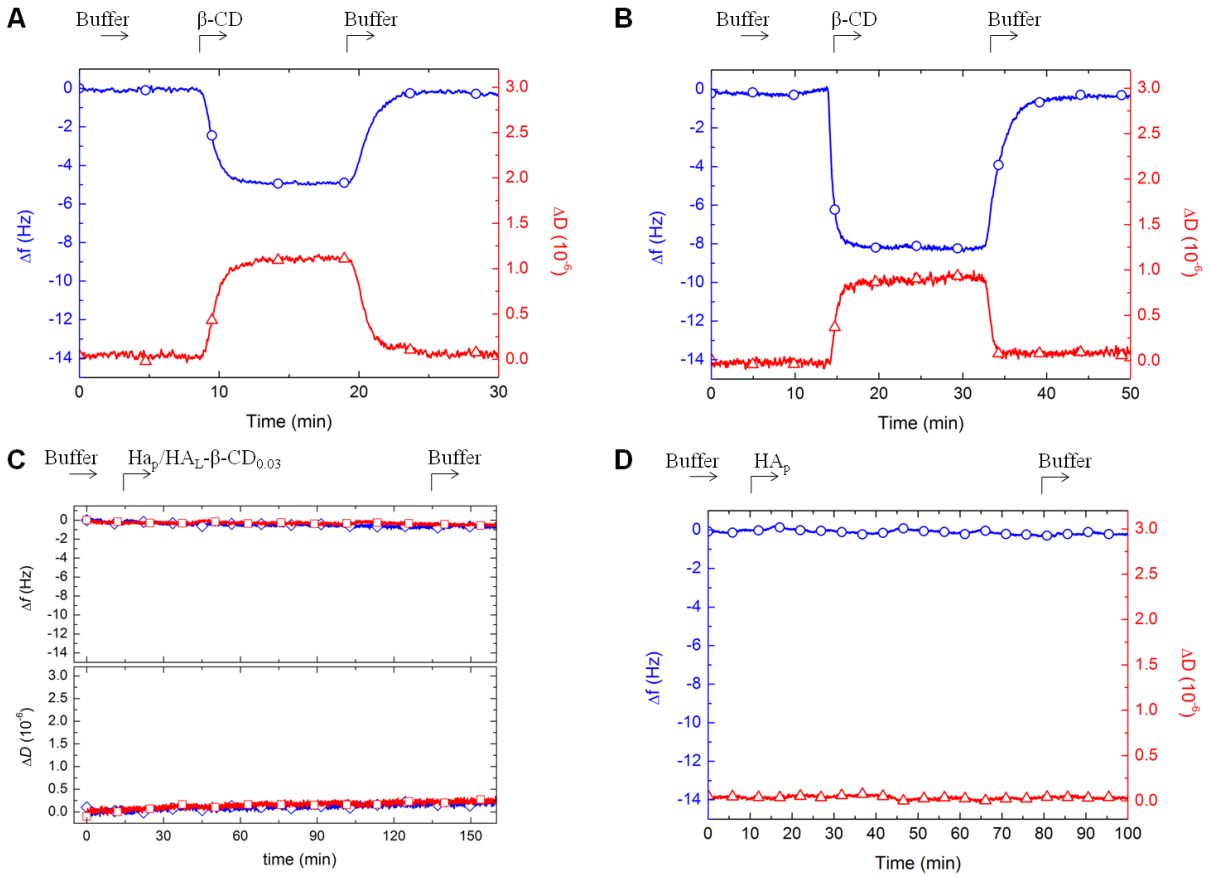


Fig. S3. Controls for specificity of binding. QCM-D response, Δf and ΔD, obtained for the binding of (A) β-CD to SAM-Fc (Γ_{guest} = 86 pmol/cm^2), (B) β-CD to SAM-AD (Γ_{guest} = 145 pmol/cm^2), (C) HA_p to SAM-Fc (Γ_{guest} = 86 pmol/cm^2, red squares) and HA_L-β-$CD_{0.03}$ to SAM (Γ_{guest} = 0 pmol/cm^2, blue lozenges), and (D) HA_p to SAM-AD (Γ_{guest} = 126 pmol/cm^2). Conditions: buffer – 10 mM HEPES (pH 7.4), 150 mM NaCl, flow rate – 20 μL/min, T = 23 °C, $c_{HA/HA\text{-}\beta\text{-}CD}$ = 120 nM, $c_{\beta\text{-}CD}$ = 5 mM.

1.3.2. Determination of HA surface density by spectroscopic ellipsometry (SE)

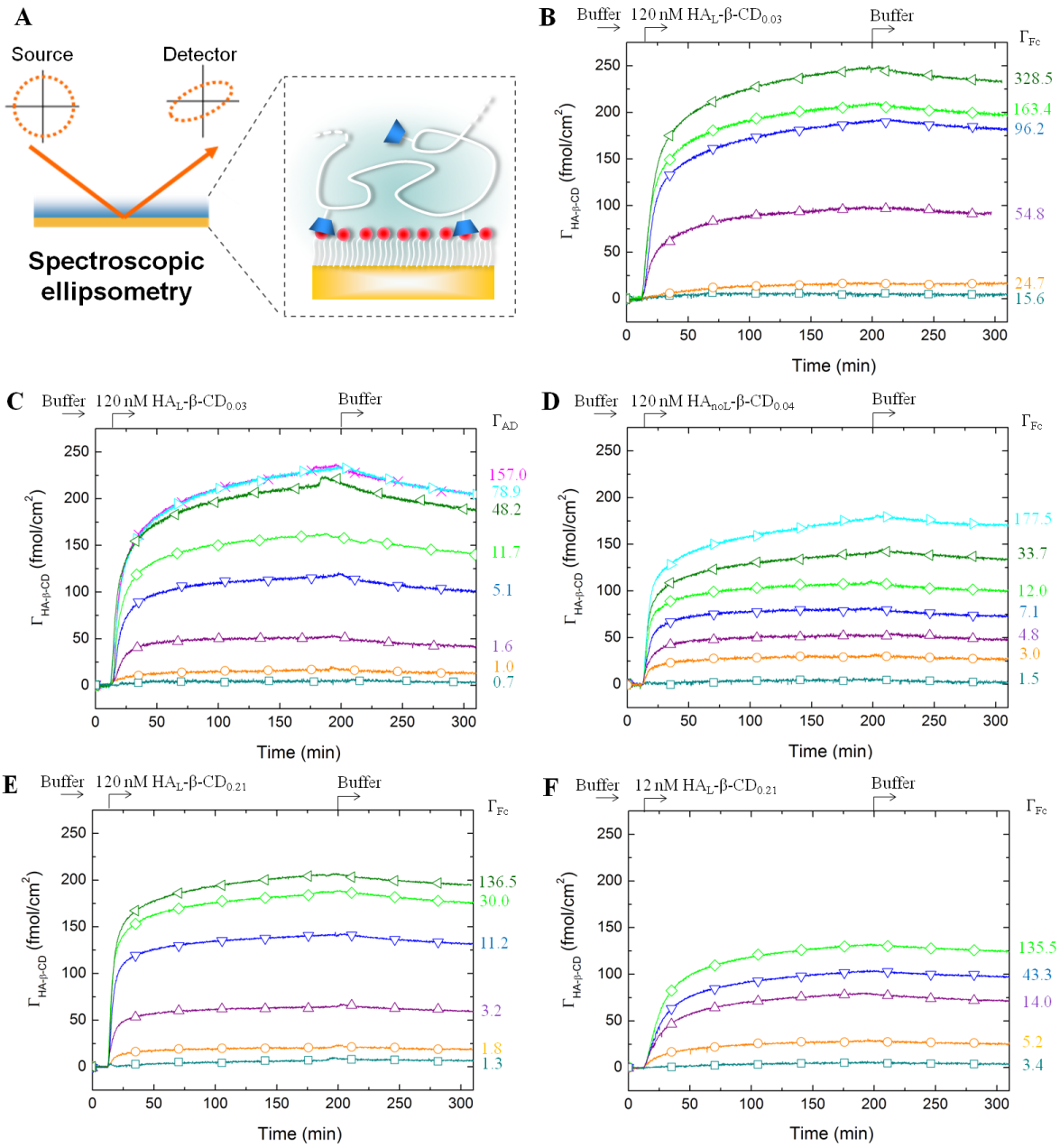


Fig. S4. HA-β-CD binding to SAMs with different guest densities monitored by SE. (A) Schematic representation of the SE setup used to characterize the surface density of the polymer layer in buffer solution. (B-F) Results obtained for the binding of HA$_L$-β-CD$_{0.03}$ to SAM-Fc (B; reproduced from ref. [2]), HA$_L$-β-CD$_{0.03}$ to SAM-AD (C), HA$_{noL}$-β-CD$_{0.04}$ to SAM-Fc (D), and HA$_L$-β-CD$_{0.21}$ to SAM-Fc (E, F). Conditions: $c_{HA-β-CD}$ = 120 nM (B-E) and 12 nM (F), buffer – 10 mM HEPES (pH 7.4) with 150 mM NaCl, T = 23 °C, flow rate = 20 µL/min. Arrows on top of the graphs indicate the start of injection with a new solution. Guest surface densities (in pmol/cm^2) for the respective measurements are indicated by numbers in matching colours on the right of the graphs.

2. Analytical model

2.1. Description of the analytical model. In this section, we briefly describe and extend an analytical model that was initially developed for reference system **(i)**.[2] The model assumes polymers to adsorb into cubic cells of size $a = R_g(4\pi/3)^{1/3}$ and allows one cell to be occupied by several polymers (R_g – polymer radius of gyration). Using this model, we analytically calculate the average number of bound polymers per cell

$$\theta(n_L) = \frac{\sum_i i\, z^i q_i}{1+\sum_i z^i q_i}, \tag{S1}$$

where the sum goes over all possible numbers i of polymers in a cell. $z \approx cN_A a^3/M_w$ is the activity of polymers in dilute solutions, with the polymer concentration c (in g/L), the polymer molecular weight M_w and Avogadro's number N_A. q_i is the bound-state partition function which counts all possible bonding arrangements between i polymers and surface ligands as

$$q_i = \frac{1}{i!}\left[\left(1 + n_L e^{-F/k_B T}\right)^{n_R} - 1\right]^i e^{-U_i/k_B T}, \tag{S2}$$

where n_L is the number of ligands per cell (guests in our experiments), n_R is the number of receptors per polymer (hosts in our experiments), T is the absolute temperature and k_B is the Boltzmann constant. F is the ligand-receptor binding free energy, which determines the probability that a particular ligand-receptor complex is formed once a polymer is in the cell. F can be estimated from the dissociation constant K_d of individual ligand-receptor interactions as

$$F = \ln(K_d a^3 N_A)\, k_B T + U_{poly} + \Delta U_{link}. \tag{S3}$$

The term U_{poly} accounts for entropic effects associated with ligand-receptor bond formation, once the polymer is present at the surface. In this model, U_{poly} represents the average entropic penalty related to the polymer backbone and the linker upon formation of a single ligand-receptor bond. To keep the model simple, we assume that polymer-bound receptors are uncorrelated within the polymer backbone, i.e. U_{poly} does not depend on the number of ligand-receptor bonds formed by a given polymer. U_{poly} is *a priori* unknown and expected to be on the order of a few $k_B T$.[7] The term ΔU_{link}, newly introduced here, takes care of entropic effects due to changes in the linker that connects the receptors to the polymer. The magnitude of ΔU_{link} is determined by the molecular details of the linker and *a priori* unknown. We quantify ΔU_{link} relative to the (arbitrarily chosen) reference system (**i**), and expect it to be on the order of a few $k_B T$.

U_i specifies the free energy penalty for the interpenetration of i polymer coils. We previously described this parameter using a simple scaling approximation

$$U_i = Ai^{9/4}, \tag{S4a}$$

for the mutual repulsion of polymers,[2] where A is an *a priori* unknown prefactor that depends on the physical properties of the polymer and that we expect to be of order $k_B T$. In reality, a polymer is also effectively repelled by the surface even in the absence of any ligand-receptor

binding, and we here refine our model by including a mean-field polymer-surface repulsion. The potential of the mean force between a self-avoiding-walk polymer and an impenetrable surface is well approximated up to 10 k_BT by an exponential function $V_{ps}(r) = 3.2\, k_BTe^{-4.17(r/R_g-0.5)}$, where r is the distance between the surface and the polymer's centre of mass.[8] The average potential of the mean force acting on a polymer when it is in the cell (i.e. the entropic penalty for the polymer to be at located at the cell surface) is then calculated through $\exp(-V_{cell}/k_BT) = \langle\exp\big(-V_{ps}(r)/k_BT\big)\rangle = \frac{1}{a}\int_0^a \exp\big(-V_{ps}(r)/k_BT\big)\,dr = 0.436$, which gives $V_{cell} = 0.83\ k_BT$. Assuming that V_{cell} stays constant for several polymers in a cell, the total free energy penalty for i polymers in a given cell due to polymer-polymer and polymer-surface repulsion becomes

$$U_i = Ai^{9/4} + V_{cell}i. \qquad \text{(S4b)}$$

Taken together, the extended model takes into account polymer-polymer and polymer-surface repulsions (through the term U_i; Eq. S4b) and the entropic penalty related to the linker and the polymer upon formation of ligand-receptor complexes (through ΔU_{link} and U_{poly}; Eq. S3). For the experimental system under study, ΔU_{link}, U_{poly} and A are the only unknown parameters.

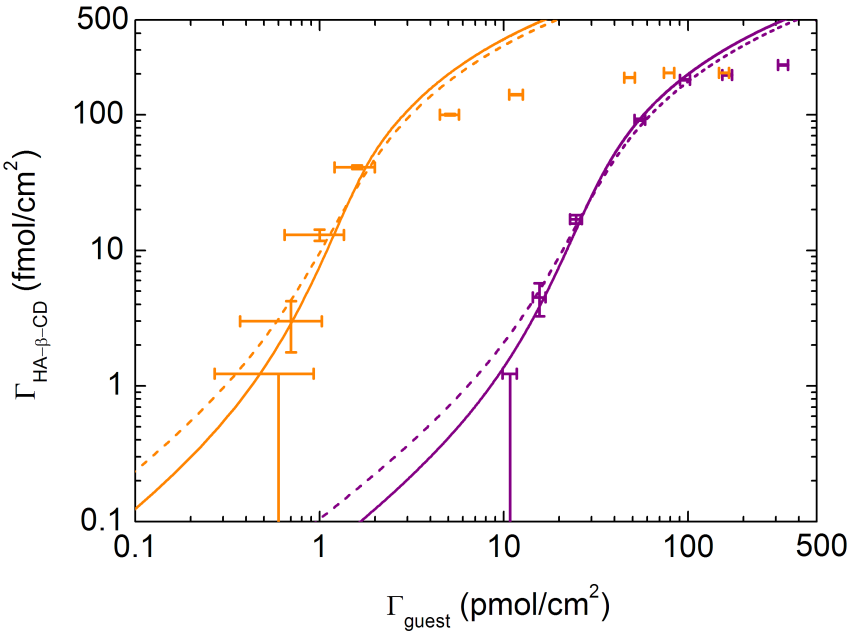


Fig. S5. Fits with the analytical models to the reference system **(i)** (purple) together with predictions for system **(iii)** (orange). The effect of the added polymer-surface repulsion in the extended model (solid lines; U_{poly} = 4.6 k_BT and A = 0.35 k_BT were fitted, ΔU_{link} = 0 was fixed) is compared to the initial model (dashed lines; U_{poly} = 4.8 k_BT and A = 0.35 k_BT were fitted, ΔU_{link} = 0 was fixed). Experimental data are from Fig. 4A and plotted in the form of error bars.

Fig. S5 illustrates that the extended model accounting for the polymer-surface repulsion reproduces experimental data more accurately at low polymer coverage than the initial model. One should, however, keep in mind that the updated model is still simplified in comparison to

the experimental system. In particular, it neglects correlations between the spatial positions of the polymer-bound receptors, uses a simple scaling approximation to describe the overlap penalty for the interpenetration of polymer coils and assumes that the entropic effects, due to constraining the polymer to the surface upon a ligand-receptor complex formation, are independent of the number of ligand-receptor complexes already present.

2.2. Derivation of a scaling parameter. In this section, we shall derive a scaling parameter from the analytical model through which the effect of modulations in the properties of multivalent polymers on superselective binding can be easily appreciated. Specifically, we consider the case of low occupancy of receptors on the polymers. A receptor can be either unbound or bound to any of the n_{L} ligands. The probability that a given receptor is bound (occupied) is $p_{\mathrm{b}} = n_{\mathrm{L}} e^{-F/k_{\mathrm{B}}T} / \left(1 + n_{\mathrm{L}} e^{-F/k_{\mathrm{B}}T}\right)$, i.e. the occupancy of receptors is low if $n_{\mathrm{L}} e^{-F/k_{\mathrm{B}}T} < 1$. We define

$$q_1 = \left(1 + n_{\mathrm{L}} e^{-F/k_{\mathrm{B}}T}\right)^{n_{\mathrm{R}}} - 1 \approx e^{n_{\mathrm{R}} n_{\mathrm{L}} e^{-F/k_{\mathrm{B}}T}} - 1, \tag{S5}$$

where we have used the approximation $e^x \approx 1 + x$ for $x < 1$ to obtain the double exponential. From Eqs. S1, S2 and S5, we find

$$\theta = \frac{\sum_i \frac{i}{i!} e^{-U_i/k_{\mathrm{B}}T} (zq_1)^i}{1 + \sum_i \frac{1}{i!} e^{-U_i/k_{\mathrm{B}}T} (zq_1)^i}. \tag{S6}$$

Eq. S6 reveals that the average number of bound polymers per cell – proportional to the polymer surface density, i.e. $\theta = \Gamma_{\mathrm{HA-\beta-CD}} N_{\mathrm{A}} a^2$ – depends exclusively on the product zq_1. With $z \approx cN_{\mathrm{A}}a^3/M_{\mathrm{w}}$, $a = R_{\mathrm{g}}(4\pi/3)^{1/3}$ and $n_{\mathrm{L}} = \Gamma_{\mathrm{guest}} N_{\mathrm{A}} a^2$, we can express

$$zq_1 \approx \frac{4\pi}{3} \frac{cN_{\mathrm{A}}R_{\mathrm{g}}^3}{M_{\mathrm{w}}} \left[\exp\left(\left(\frac{3}{4\pi}\right)^{1/3} \frac{n_{\mathrm{R}}\Gamma_{\mathrm{guest}}}{K_{\mathrm{d}}R_{\mathrm{g}}} e^{-(\Delta U_{\mathrm{link}} + U_{\mathrm{poly}})/k_{\mathrm{B}}T}\right) - 1\right] \tag{S7}$$

in terms of experimental parameters. The derivation shows that, with polymer size R_{g} and molar concentration c/M_{w} maintained constant, θ depends exclusively on the scaling parameter

$$x_{\mathrm{S}} = \frac{n_{\mathrm{R}}\Gamma_{\mathrm{guest}}}{K_{\mathrm{d}}/\rho_0} e^{-(\Delta U_{\mathrm{link}} + U_{\mathrm{poly}})/k_{\mathrm{B}}T}. \tag{S8}$$

We have added the standard concentration $\rho_0 = 1$ M to x_{S}, to obtain a parameter that has the units of Γ_{guest}.

Fig. 4C demonstrates that all experimental data sets for a given polymer concentration, when properly re-scaled, essentially fall on a single master curve. This identifies the scaling variable as a simple tool to quantitatively predict how the properties of a multivalent polymer can be tuned to target a desired superselectivity range. For completeness, Fig. S6 illustrates how polymer size and concentration affect the quality and the range of superselectivity.

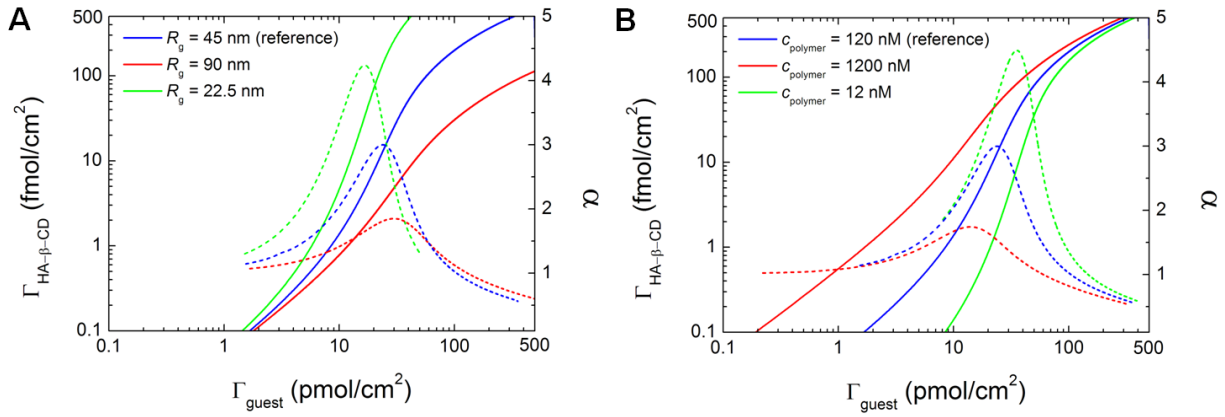


Fig. S6. Dependence of superselective binding on polymer radius of gyration and concentration. Predictions with the analytical model for $\Gamma_{HA\text{-}\beta\text{-}CD}$ *vs* Γ_{guest} dependences (solid lines) at a 2-fold reduced/increased R_g (A) and at a 10-fold reduced/increased c (B), with all other parameters kept identical to the best fit for reference system (**i**). Selectivity parameter α as a function of Γ_{guest} for the model data (dotted lines). The graphs show that the quality of superselectivity, or magnitude of α, increases with decreasing polymer concentration and size. The ability of multivalent polymers to bind superselectively is due to the sharp increase of q_i with n_L (Eq. S2). However, the formula for $\theta(n_L)$ has q_i in both the numerator and denominator (Eq. S1), and θ can benefit from the rapid variation of q_i only when $z^i q_i \ll 1$. One can maintain $z^i q_i \ll 1$ through reducing z, which can be achieved by decreasing c/M_w or R_g.

3. Numerical simulations

3.1 Soft-blob model. The simulations were performed with a soft-blob model for polymers presented in detail elsewhere.[9] Briefly, each polymer chain is represented by N_b soft Gaussian blobs with radius of gyration r_b that are connected *via* harmonic springs

$$U_{ch} = 0.534 k_B T (r/r_b - 0.730)^2, \quad \text{(S9)}$$

where r is the centre-to-centre distance of adjacent blobs. The blob-blob interaction is described through a Gaussian repulsion

$$U_{bb} = 1.75 k_B T e^{-0.80(r/r_b)^2}, \quad \text{(S10)}$$

while the blob-surface interaction is modeled as an exponential repulsion

$$U_{bs} = 3.20 k_B T e^{-4.17(r/r_b - 0.50)} \quad \text{(S11)}$$

A convenient feature of the soft-blob model are transferable potentials, i.e. one can represent a given polymer by many small blobs, by a few larger ones or by a single large blob, while keeping the interaction potentials the same. The radius of gyration of the polymer is given by

$$R_g = r_b N_b^{0.588}. \quad \text{(S12)}$$

We assume that each individual blob represents a polymer in the scaling regime, hence r_b must be large enough to contain at least a few statistical Kuhn segments. Moreover, the model considers pair wise but not higher order interactions, i.e. it is appropriate for studying dilute and semi-dilute polymer solutions, where the blob density does not exceed 1 blob per blob volume $v_\mathrm{b} < 3/(4\pi r_\mathrm{b}^3)$. In our simulations of HA-β-CD binding to guest-functionalized surfaces, we found this to be the case for $N_\mathrm{b} \geq 20$.

We assume that there are n_R receptors per polymer (hosts in our experiments) and that these are randomly distributed along the chain. Blobs can therefore carry zero, one or several receptors. Ligands (guests in our experiments) are modelled as immobile point-like objects that are randomly distributed on the impenetrable and planar surface. Receptor-carrying blobs can bind to the surface-attached ligands if the blob's centre of mass is within the distance r_b from the ligand. The binding is specific (i.e. there are no interactions between blobs and ligands other than ligand-receptor binding) and valence limited (i.e. one ligand can bind only a single receptor and *vice versa*). Typical simulation snapshots obtained for soft-blob polymers adsorbed to surfaces with different ligand densities are shown in Fig. S7.

The ligand-receptor binding free energy (F) reflects pure ligand-receptor interaction, but it also depends on r_b and on the linker properties. We assume that the unbound receptor can explore the space within the blob and that different receptors in a blob are uncorrelated. We use the same arguments to calculate the single ligand-receptor binding free energy as in the analytical theory (Eq. S3), except that now we apply it to an individual blob within a polymer

$$F_\mathrm{sim} = \ln(K_\mathrm{d}\,\tfrac{4\pi}{3} r_\mathrm{b}^3\,N_\mathrm{A})\,k_\mathrm{B}T + U_\mathrm{poly,b} \qquad \text{(S13)}$$

We note that $U_\mathrm{poly,b}$ is distinct from U_poly in Eq. S3 in that entropic penalties related to generic polymer properties are implicitly considered in the simulations down to the size of the blob and thus no longer contained in $U_\mathrm{poly,b}$.

3.2 Monte Carlo algorithm. To determine the equilibrium number of surface-bound polymers, we perform grand canonical Monte Carlo (MC) simulations using the soft-blob model. We simulate a box of size $L_x \times L_y \times L_z$ with periodic boundary conditions in the orthogonal directions x and y along the surface. The impenetrable planar surface containing immobile ligands is positioned at $z = 0$ (where z denotes the coordinate normal to the surface). The box height is large enough ($L_z > 3R_\mathrm{g}$) such that bound polymers are not constrained by the box ceiling. Typical lateral dimensions are $L_x = L_y \approx 5\ R_\mathrm{g}$. The box is in contact with a polymer solution with a fixed chemical potential μ that is chosen to reproduce polymer concentrations in the bulk solutions specific to our experiments.

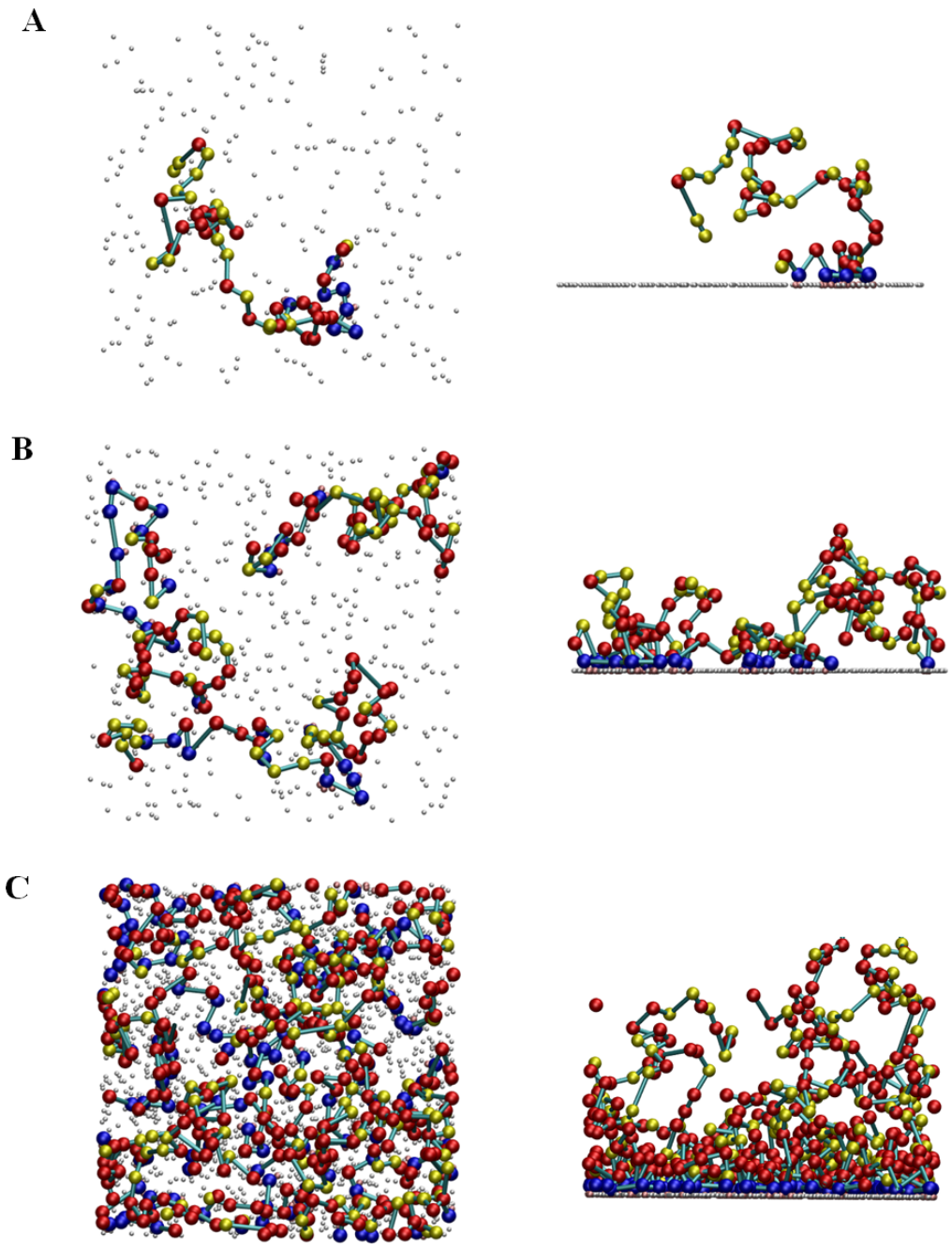


Fig. S7. Snapshots (left – top view, right – side view) of Monte Carlo simulations of soft-blob polymers bearing receptors at ligand-coated surfaces. Unbound ligands are represented as grey spheres, and bound ligands as pink spheres. The polymers are represented by blobs that carry either no receptor (red), exclusively unbound receptors (yellow) or at least one bound receptor (blue). The blobs are connected into chains with turquoise rods. Conditions: $N_b = 50$, $F = -3\ k_BT$ and different numbers of ligands: 244 ligands (2.21 pmol/cm^2), 1 polymer chain (A), 403 ligands (3.66 pmol/cm^2), 3 polymer chains (B) and 1808 ligands (16.4 pmol/cm^2), 12 polymer chains (C); the system size in lateral directions is 3 R_g.

A particular state in our simulations is determined by the vector positions of all blobs and the arrangement of ligand-receptor bonds. For the efficient sampling of the states, we employ two types of MC moves: (i) polymer insertion/deletion moves and (ii) single blob translational moves that are integrated with ligand-receptor binding. The polymers are generated or deleted using Rosenbluth sampling with configurational bias.[10] Only non-bound polymers can be inserted or deleted. Each simulation starts with an empty box and finishes after ~10^{11} MC cycles. In each cycle, we randomly select either to insert or delete a polymer (with probability $p_{\mathrm{ins/del}} = 1(N_{\mathrm{b}} + 1)$), or to move a single blob (with probability $p_{\mathrm{hop}} = N_{\mathrm{b}}(N_{\mathrm{b}} + 1)$). After each step, the number of bound polymers (i.e. polymers attached to the surface *via* at least one ligand-receptor bond) is determined.

In order to speed up the simulations, the binding between ligands and receptors is integrated within the blob translational moves. The binding partition function which counts all possible ways to make m_{b} bonds within a blob is calculated as

$$q_{m_{\mathrm{b}}} = e^{-m_{\mathrm{b}}F/k_{\mathrm{B}}T}\,\frac{n_{\mathrm{R,b}}!n_{\mathrm{L,f}}!}{(n_{\mathrm{R,b}}-m)!m!(n_{\mathrm{L,f}}-m)!}, \tag{S14}$$

where we have $n_{\mathrm{L,f}}$ free ligands within reach of the $n_{\mathrm{R,b}}$ receptors of the blob and consider all possible ways of binding together with m_{b} bonds. The total binding partition function considers all possible numbers of bonds

$$q_{\mathrm{b}} = \sum_{m_b=0}^{\min\,(n_{\mathrm{R,b}},n_{\mathrm{L,f}})} q_{m_{\mathrm{b}}}. \tag{S15}$$

The above expression is essentially the same as the exact binding partition function used in the analytical model. A new trial position for a given blob is then considered and its new internal binding partition function $q_{\mathrm{b}}^{\mathrm{n}}$ is calculated. The move is then accepted with probability

$$p_{\mathrm{o}\to\mathrm{n}} = \min\left[1, \frac{q_{\mathrm{b}}^{\mathrm{n}}}{q_{\mathrm{b}}^{\mathrm{o}}} e^{(U_{\mathrm{o}}-U_{\mathrm{n}})/k_{\mathrm{B}}T}\right], \tag{S16}$$

where the indices "o" and "n" indicate the old and new positions, respectively, and U is the potential energy of the system determined by Eqs. S9-11.

Regardless of whether the move is accepted or not, the blob of interest will still have all receptors unbound. We now randomly choose how many bonds to form. The probability to form m_{b} bonds is

$$p_{m_{\mathrm{b}}} = \frac{q_{m_{\mathrm{b}}}}{q_{\mathrm{b}}} \tag{S17}$$

where we use the partition functions calculated for the new/old blob position (Eqs. S14-15) if the translation move was accepted/not-accepted. To form m_{b}^* bonds, we randomly chose m_{b}^* ligands and m_{b}^* receptors and randomly bind them together. The probability to choose a particular state (bonding arrangement) with m_{b}^* bonds is then

$$p_{\mathrm{s}} = \frac{e^{-m_{\mathrm{b}}^* F/k_{\mathrm{B}}T}}{q_{\mathrm{b}}}. \tag{S18}$$

The equilibrium probability for the system to be in the state (1), with potential energy U_1 and number of ligand-receptor bonds m_1 is proportional to the total free energy of that state

$$p_1 \propto e^{-(U_1+m_{\mathrm{b},1}F)/k_BT}. \tag{S19}$$

The transition probability going from state (1) to state (2) (with potential energy U_2 and $m_{\mathrm{b},2}$ bonds) is

$$\prod_{1\to2} = \min\left[1, \frac{q_{\mathrm{b}2}}{q_{\mathrm{b}1}} e^{(U_1-U_2)/k_BT}\right] \times \frac{e^{-m_{\mathrm{b},2}F/k_{\mathrm{B}}T}}{q_{\mathrm{b}2}}, \tag{S20}$$

where the first factor corresponds to the probability of moving the blob to a new position (Eq. S16) which is multiplied by the probability that we choose a particular bond arrangement at the new position (Eq. S18). The reverse transition probability is

$$\prod_{2\to1} = \min\left[1, \frac{q_{\mathrm{b}1}}{q_{\mathrm{b}2}} e^{(U_2-U_1)/k_{\mathrm{B}}T}\right] \times \frac{e^{-m_{\mathrm{b},1}F/k_{\mathrm{B}}T}}{q_{\mathrm{b}1}}, \tag{S21}$$

and the ratio between the two is obtained, given that $\frac{q_{b2}}{q_{b1}} e^{(U_1-U_2)/k_BT}$ can be either greater or smaller than 1

$$\frac{\prod 1\to2}{\prod 2\to1} = \frac{e^{-(U_2+m_{\mathrm{b},2}F)/k_{\mathrm{B}}T}}{e^{-(U_1+m_{\mathrm{b},1}F)/k_{\mathrm{B}}T}} = \frac{p_2}{p_1}. \tag{S22}$$

Both cases give the same result, proving that the algorithm obeys detailed balance. The method was also tested to provide the same equilibrium result when compared to standard bind/unbind MC moves.

Each simulation was initiated with an empty box and ran for ~10^{11} MC cycles. The equilibrium number of polymers was determined by averaging over the number of adsorbed polymers in the second half of the simulation run. We tested for convergence by considering the average number of polymers in intervals of 10^{10} MC cycles. Simulations were not considered converged as long as this number was increasing over subsequent intervals; when the averaged number of polymers started fluctuating, we considered the simulations sufficiently equilibrated. We also monitored the relaxation of the average number of bonds per polymer, which is related to the brush conformation. Simulations with N_b=20 blobs per polymer reached equilibrium. Using longer polymers (N_b=50) simulations at the largest polymer coverages did not reach equilibrium within 2×10^{11} MC cycles. However, simulations did converge for lower polymer coverages ($\Gamma_{\text{HA-β-CD}} < 100$ fmol/cm^2) and the snapshots shown in Fig. S7 are representative of an equilibrium configuration.

3.3. Dependence of U_{poly} on m. To determine how U_{poly} depends on the number of bonds m, we calculated the free energy of binding directly from simulations, using the Wang-Landau technique,[11] and then compared the results to the analytical model. To this end, simulations were performed with a single, surface-constrained polymer chain for $0 \le m \le n_{\mathrm{R}}$ bonds. For simplicity, the calculation of the free energy F_{WL} was performed at a receptor-ligand binding energy within a blob $\epsilon_{\mathrm{WL}} = 0$; this does not limit the generality of the approach (i.e. the

receptor-ligand energy can be added later). The reference state was chosen to be the unbound polymer, i.e. $F_{\mathrm{WL}}(m=0)=0$, and the polymer's centre of mass was constrained to lie within a distance $a=(4\pi/3)^{1/3}R_{\mathrm{g}}$ from the surface, to have a consistent comparison with the analytical model (where we assumed the polymer to be located in a cell of size *a*).

The F_{WL} obtained from simulations is related to the analytical theory through

$$e^{-[F_{\mathrm{WL}}(m)+\epsilon(r_{\mathrm{b}})m]/k_{\mathrm{B}}T}=\binom{n_{\mathrm{L}}}{m}\binom{n_{\mathrm{R}}}{m}m!\,e^{-m[U_{\mathrm{poly,eff}}(m)+\epsilon(R_{\mathrm{g}})]/k_{\mathrm{B}}T}. \qquad \text{(S23)}$$

The term on the left relates to the simulations, where the receptor-ligand binding energy within a blob of size r_{b}, i.e. $\epsilon(r_{\mathrm{b}})=\ln(K_{\mathrm{d}}(4\pi r_{\mathrm{b}}/3)^3N_{\mathrm{A}})$, is now added. The term on the right is the full partition function for a single polymer of the analytical theory,[2] where the combinatorial factor explicitly considers all possible distinct ways of binding together n_{L} ligands with n_{R} receptors through m bonds and where $\epsilon(R_{\mathrm{g}})=\ln(K_{\mathrm{d}}a^3N_{\mathrm{A}})$. The subscript 'eff' indicates that the simulations consider only effects that reflect generic polymer properties. In reality, there will be an additional term related to effects that are not explicitly considered in the simulations, such as the effect of the polymer linker. However, these effects should not depend on *m*, i.e. $U_{\mathrm{poly}}(m)=U_{\mathrm{poly,eff}}(m)+\mathrm{const}$ and the dependence of U_{poly} on *m* is thus fully contained in $U_{\mathrm{poly,eff}}(m)$. Re-writing the above equality, using $R_{\mathrm{g}}=r_{\mathrm{b}}N_{\mathrm{b}}^{\upsilon}$ with the scaling exponent $\upsilon=0.588$, we obtain

$$U_{\mathrm{poly,eff}}(m)/k_{\mathrm{B}}T=\frac{1}{m}F_{\mathrm{WL}}(m)/k_{\mathrm{B}}T+\frac{1}{m}\ln\left(\frac{n_{\mathrm{L}}!n_{\mathrm{R}}!}{(n_{\mathrm{L}}-m)(n_{\mathrm{R}}-m)!m!}\right)-3\upsilon\ln(N_{\mathrm{b}}). \qquad \text{(S24)}$$

$U_{\mathrm{poly,eff}}$ as a function of *m* is shown in Fig. S8 for the experimentally relevant polymer valencies n_{R}. One can see that $U_{\mathrm{poly,eff}}$ is not constant, as we assumed in the analytical model, but instead decreases monotonically with *m*. This implies that the individual bonds effectively strengthen with *m* and that the formation of additional bonds is facilitated. This cooperative effect increases the quality of superselectivity of multivalent polymer binding and thus explains why the simulations predict a somewhat higher quality of superselectivity than the analytical model (Fig. 4). Furthermore, one can see that $U_{\mathrm{poly,eff}}$ weakly decreases with n_{R} for a given *m*. This explains why the analytical model underestimated the shift in the superselectivity range experimentally observed when switching from low to high polymer valency. Indeed, the differences in $U_{\mathrm{poly,eff}}$ between low and high valency polymers for large enough *m* (Fig. S8) are comparable to the discrepancy of 0.7 $k_{\mathrm{B}}T$ that we had found between experiment and analytical model.

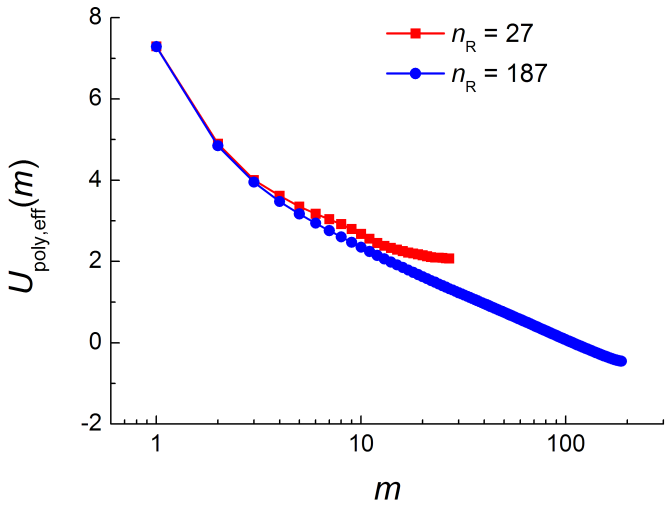


Fig. S8. $U_{poly,eff}$ as a function of the number of bonds m obtained from simulations at N_b=100 and n_R = 27 and 187. The free energy calculations were averaged over at least 100 different random realizations of the receptor positions on a chain; standard deviations are smaller than the symbol sizes.

Supporting references

(1) Macosko, C. *Rheology: Principles, Measurements and Applications*; Wiley-VCH: New York, 1994.

(2) Dubacheva, G. V.; Curk, T.; Mognetti, B. M.; Auzély-Velty, R.; Frenkel, D.; Richter, R. P. *J. Am. Chem. Soc.* **2014**, *136*, 1722.

(3) Johannsmann, D. *Macromol. Chem. Phys.* **1999**, *200*, 501−516.

(4) Johannsmann, D. http://www2.pc.tu-clausthal.de/dj/software_en.shtml.

(5) Eisele, N. B.; Andersson, F. I.; Frey, S.; Richter, R. P. *Biomacromolecules* **2012**, *13*, 2322.

(6) Rubinstein, M.; Colby, R. H. *Polymer Physics*; Oxford University Press: Oxford, 2003.

(7) Martinez-Veracoechea, F. J.; Leunissen, M. E. *Soft Matter* **2013**, *9*, 3213.

(8) Martinez-Veracoechea, F. J.; Bozorgui, B.; Frenkel, D. *Soft Matter* **2010**, *6*, 6136.

(9) Pierleoni, C.; Capone, B.; Hansen, J.-P. *J. Chem. Phys.* **2007**, *127*.

(10) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*; Academic Press, Inc., 1996.

(11) Wang, F.; Landau, D. P. *Phys. Rev. Lett.* **2001**, *86*, 2050.