bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

A quantitative view of strategies to engineer cell-selective ligand binding

This manuscript was automatically generated on November 25, 2020.

Authors

- **Zhixin Cyrillus Tan**
  ORCID 0000-0002-5498-5509 · Github cyrillustan
  Bioinformatics Interdepartmental Program, University of California, Los Angeles; These authors contributed equally.

- **Brian Orcutt-Jahns**
  ORCID 0000-0002-1436-1224 · Github borcuttjahns
  Department of Bioengineering, University of California, Los Angeles; These authors contributed equally.

- **Aaron S. Meyer**
  ORCID 0000-0003-4513-1840 · Github aarmey · twitter aarmey
  Department of Bioengineering, University of California, Los Angeles; Bioinformatics Interdepartmental Program, University of California, Los Angeles; Jonsson Comprehensive Cancer Center, University of California, Los Angeles; Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, University of California, Los Angeles

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

A critical property of many therapies is their selective binding to specific target populations. Exceptional specificity can arise from high-affinity binding to unique cell surface targets. In many cases, however, therapeutic targets are only expressed at subtly different levels relative to off-target cells. More complex binding strategies have been developed to overcome this limitation, including multi-specific and multi-valent molecules, but these create a combinatorial explosion of design possibilities. Therefore, guiding strategies for developing cell-specific binding are critical to employ these tools. Here, we extend a multi-valent binding model to multi-ligand and multi-receptor interactions. Using this model, we explore a series of mechanisms to engineer cell selectivity, including mixtures of molecules, affinity adjustments, and valency changes. Each of these strategies maximizes selectivity in distinct cases, leading to synergistic improvements when used in combination. Finally, we identify situations in which selectivity cannot be derived through passive binding alone to highlight areas in need of new developments. In total, this work uses a quantitative model to unify a comprehensive set of design guidelines for engineering cell-specific therapies.

Summary points

- Affinity, valency, and other alterations to target cell binding provide enhanced selectivity in specific situations.
- Evidence for the effectiveness and limitations of each strategy are abundant within the drug development literature.
- Combining strategies can offer enhanced selectivity.
- A simple, multivalent ligand-receptor binding model can help to direct therapeutic engineering.

Introduction

Many drugs both derive their therapeutic benefit and avoid toxicity through selective binding to specific cells within the body. Often, target cells can differ from off-target populations only subtly in surface receptor expression, making selective activation of target cells difficult to achieve. Even with a drug of very specific molecular binding, genetic and non-genetic heterogeneity can create a wide distribution of cell responses. This can result in reduced effectiveness and increased toxicity. Specific cell targeting is a universal challenge in protein-based therapies. For example, in cancer, resistance to anti-tumor antibodies,[1] targeted inhibitors,[2] chemotherapies,[3] and chimeric antigen receptor T cells[4,5] all can arise through the selection of target cells among heterogeneous cell populations. While the immune system takes advantage of heterogeneity at the single-cell level to translate noisy inflammatory signals into robust yet sensitive responses,[6] this heterogeneity impedes our effort of creating a highly specific drug. The intricacies of both inter-population receptor expression differences and intra-population receptor expression heterogeneity present significant challenges that limit the selectivity of therapies within the body.

Further improving cell-specific targeting will require new strategies of engineering specificity. Non-cellular therapies such as protein therapies have most extensively been engineered to target specific cell types through mutations that provide high-affinity binding to unique surface antigens.[7] This

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

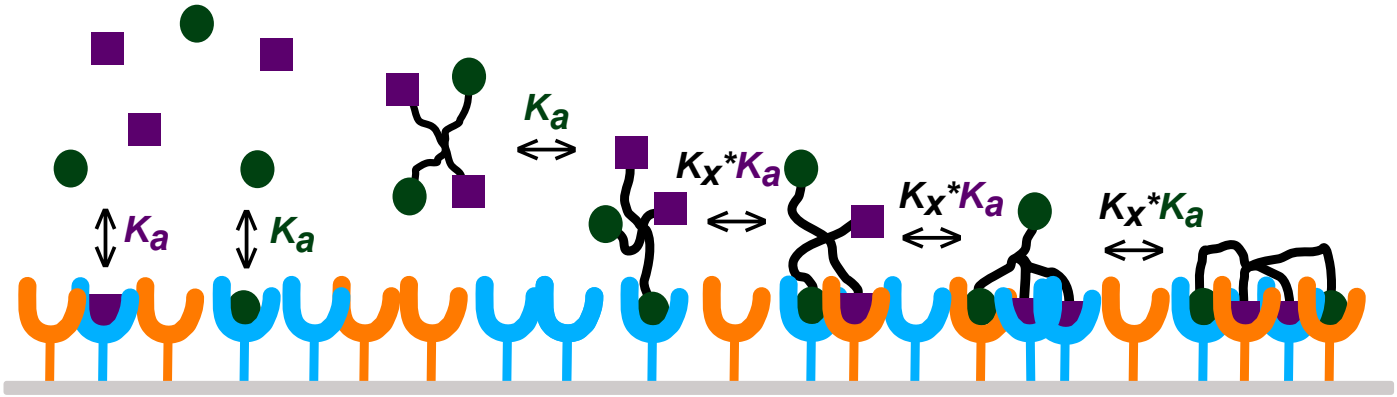
strategy can enhance specificity, but only to a limited degree, particularly when target cells can only be distinguished by subtle quantitative differences in surface antigen or by combinations of markers. The limitations of single-antigen targeting have led to efforts to program complex logic into cellular therapies to recognize target cells more specifically.[8–10] However, non-cellular therapies have considerable benefits in drug access, manufacturing, and reliability; some of the same benefits have begun to be engineered into these agents.

The enormous number of potential configurations of ligand designs make computational tools essential for designing highly selective therapies.[11] Here, we analyze a suite of molecular approaches for engineering cell-specific binding using a multi-valent, multi-receptor, multi-ligand binding model. We show that strategies including affinity, valency, binding competition, ligand mixtures, and hetero-valent complexes provide distinct improvements in cell-specific targeting. Finally, we combine these strategies to target cells through combinatorial methods. In total, our results demonstrate that multi-valent ligands can offer offer effective cell specific targeting without the need to engineer complex cellular therapies, and that their design can be guided using a mechanistic binding model.

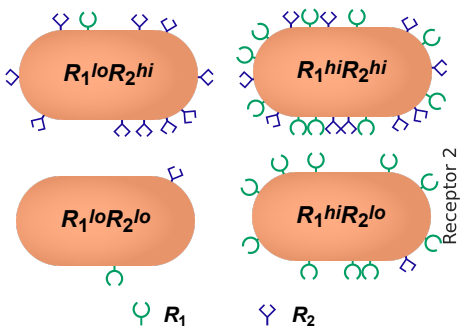
Results

A Model System to Explore the Factors Contributing to Cell Selectivity

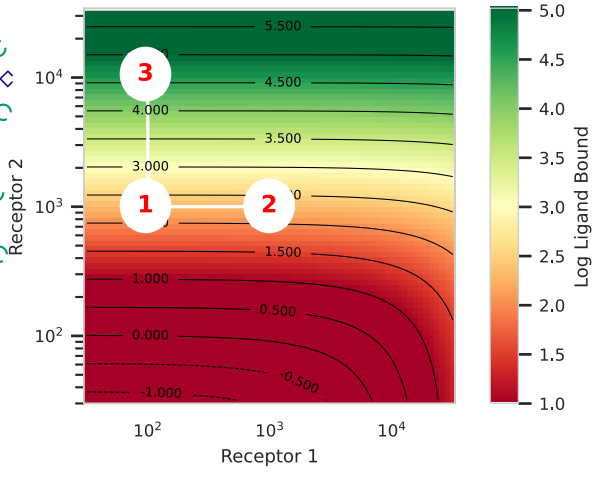
a



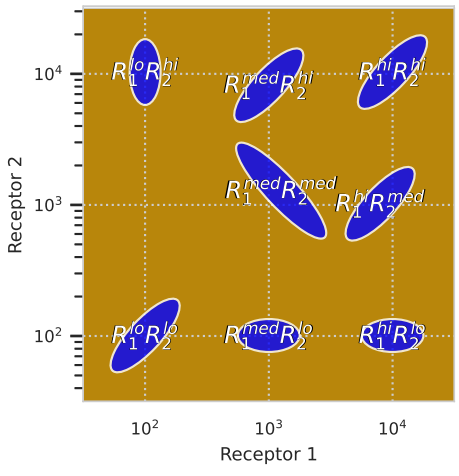
b



c



d



bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 1: A model system for exploring the factors contributing to cell selectivity. a) A simplified schematic of the binding model. There are two types of receptors and two types of ligand monomers that form a tetravalent complex. b) A cartoon for four cell populations expressing two different receptors at low or high amounts. c) A sample heat/contour map for the model-predicted log ligand bound given the expression of two types of receptors. d) Eight arbitrary theoretical cell populations with various receptor expression profiles.

Here we investigate cell-specific targeting quantitatively by extending a multi-valent, multi-ligand equilibrium binding model. Virtually any therapy, including monoclonal antibodies, small-molecule inhibitors, and cytokines, can be thought of as ligands for respective receptors expressed on target cells. Ligand binding to target cells is a necessary first step, and essential for a drug's intended action. In contrast, binding to off-target cells can result in unintended effects or toxicity. Some cell populations can be distinguished by their expression of a unique receptor not expressed by other populations, but more commonly, target and off-target cells express the same collection of receptors and differ only in their magnitudes of receptor expression. In such situations, engineering drugs to optimize target cell binding while minimizing their binding to off-target cells is an area of ongoing research, and has inspired a myriad of drug design strategies.[12–14] In this work, we define cell population selectivity as the ratio of the number of ligands bound to target cell populations divided by the number of ligands bound to off-target cell populations. We will use a quantitative binding estimation for each cell population to examine these strategies.

While ligand-receptor binding events in biology are diverse, they are governed by thermodynamic properties and the law of mass action. For the binding between monomer ligands and receptors, their affinity can be described by the association constant K_a, or its reciprocal, the dissociation constant K_d. Binding behavior is more complicated when the ligands are multivalent complexes consisting of multiple units, each of which can bind to a receptor (Fig. 1a). During initial association, we assume that the first subunit on a ligand complex binds according to the same dynamics that govern monovalent binding. Subsequent binding events exhibit different behavior, however, due to the increased local concentration of the complex and steric effects. Here, we assume that the effective association constant for the subsequent bindings is proportional to that of the free binding, but scaled by a crosslinking constant, K_x^*, which describes how easily a multivalent ligand bound to a cell monovalently can attain secondary binding. Another consideration that must be made when modeling multivalent binding processes is whether ligand complexes are formed via random assortment of monomers, or whether the monomer composition is uniform across complexes by engineering. We developed a multivalent binding model that calculates the amount of ligand bound at equilibrium taking each of these factors into account (see methods).

As a simplification, we will consider theoretical cell populations that express only two receptors capable of binding ligand (Fig. 1b), ranging in abundance from 100 to 10,000 per cell. Figure 1c shows the log-scaled predicted amount of binding of a monovalent ligand given the abundance of two receptors, with the concentration of ligand, L_0, to be $1\,\mathrm{nM}$, and its dissociation constants to the two receptors to be $10\,\mu\mathrm{M}$ and $100\,\mathrm{nM}$, respectively. Because all axes are log-scaled, the number of contour lines between two points indicates the ratio of ligand binding between populations. For instance, in figure 1c, cell populations at points 1 and 2 are on the same contour line and thus have the same amount of ligand bound; the cell populations at points 1 and 3 are separated by multiple contour lines, indicating that cells at point 3 bind more ligands (In fact, the ratio can be read as the exponent of the contour line difference. For point 3 to point 1, the ratio is $e^{4.6-2.3} \approx 7.4$). Alternatively, we can think of moving from

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

one point to another as a change of expression profile for a cell population. This situation might correspond to some cues inducing expression of a receptor, such as interferon-induced upregulation of MHC and regulatory T cell upregulation of IL-2Rα.[15] When the amount of receptor 1 (R_1) increases on a cell (a point moves rightward, e.g. from 1 to 2), the amount of binding doesn't increase significantly. On the contrary, a cell with increased expression of R_2 (moving upward, e.g. from 1 to 3) will bind significantly more ligands. Here, the ligand's high affinity for R_2 and relatively low affinity for R_1 lead to binding varying more strongly with changes to R_2 expression than R_1.

To analyze more general cases, we arbitrarily defined eight theoretical cell populations according to their differential expression of two receptor types (R_1 and R_2 plotted on x and y axes). As shown in Fig 1d, they either have high (10^4), medium (10^3), or low (10^2) expression of R_1 and R_2. The receptor expression profile within each cell population can also vary widely. To demonstrate cell-to-cell heterogeneity, we also defined intrapopulation variability of each population arbitrarily. For instance, the expression profile of $R_1^{med}R_2^{med}$ has a wider range. The intrapopulation variance will be accounted for using sigma point filters.[16] We will use this binding model to examine how engineering a ligand using various strategies can improve cell-specific targeting. Although we will only consider two receptor and ligand subunit types respectively, the principles we present can generalize to more complex cases.

Affinity Provides Selectivity Toward Cell Populations with Divergent Receptor Expression

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

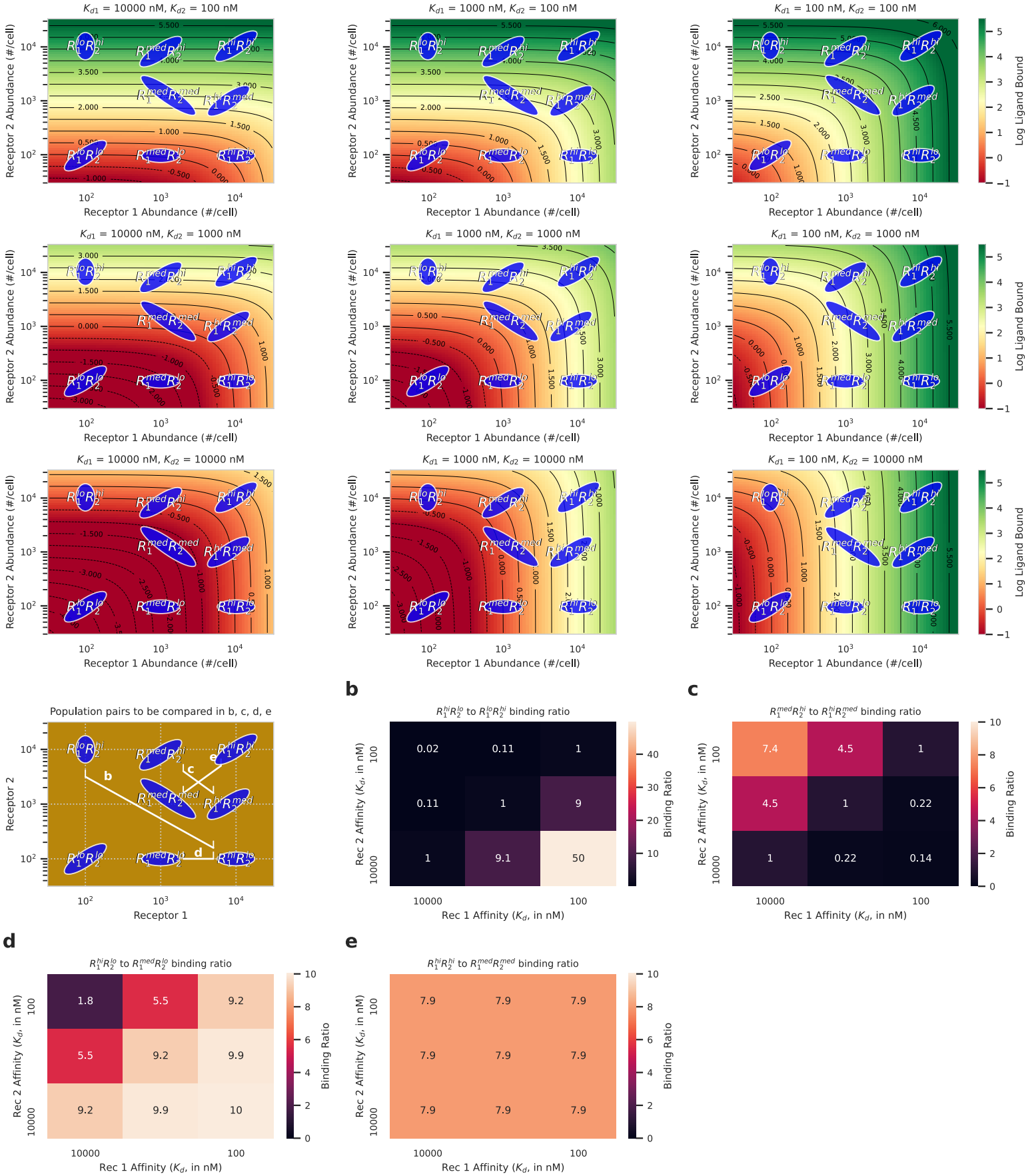


Figure 2: Affinity provides selectivity to cell populations with divergent receptor expression. a) Heat/contour maps of monovalent ligand binding to cell populations given the surface abundance of two receptors. Ligand dissociation constants to these receptors range from $10\mu\text{M} \sim 100\text{nM}$. Ligand concentration $L_0 = 1\text{nM}$. b-e) Heatmap of binding ratio of cell

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC BY 4.0 International license.

populations exposed to a monovalent ligand with dissociation constants to receptor 1 and 2 ranging from $10^4 \sim 10^2 \mathrm{nM}$, at a concentration $L_0 = 1\mathrm{nM}$. Ligand bound ratio of (b) $R_1^{hi}R_2^{lo}$ to $R_1^{lo}R_2^{hi}$, (c) $R_1^{med}R_2^{hi}$ to $R_1^{hi}R_2^{med}$, (d) $R_1^{hi}R_2^{lo}$ to $R_1^{med}R_2^{lo}$, and (e) $R_1^{hi}R_2^{hi}$ to $R_1^{med}R_2^{med}$.

We first explored altering the ligand binding affinity as an engineering strategy to enhance its cell population specificity. Here, we showed the binding pattern of monovalent ligands with various affinities ranging from $10\ \mu\mathrm{M}$ to $100\ \mathrm{nM}$ to both receptors R_1 and R_2 (Fig. 2a). We found that when a target cell population expresses a receptor not expressed by off-target cell populations, enhancing the affinity of the drug to this receptor is a clear and effective strategy to increase selective binding to this population. For example, $R_1^{hi}R_2^{lo}$ only significantly expresses R_1, while $R_1^{lo}R_2^{hi}$ only significantly expresses R_2. When the affinity to R_1 is enhanced and the affinity to R_2 is reduced, the binding selectivity towards $R_1^{hi}R_2^{lo}$ increases (Fig. 2b). The contour plots in Fig. 2a shows this trend more intuitively: when ligand affinity to R_1 increases, which corresponds to shifting from subplots on the left to the ones on the right, binding is shown to vary more strongly according to R_1 expression, indicating an increase in the amount of ligand bound for the populations with high R_1 expression, such as $R_1^{hi}R_2^{hi}$, $R_1^{hi}R_2^{med}$, and $R_1^{hi}R_2^{lo}$.

However, situations where both on- and off-target cell populations express the same set of receptors and differ only in their magnitude of expression are just as common. In these cases, we found out that it is beneficial for the drug to bind tightly to the most comparatively highly expressed receptor on the target population. Some examples of this pattern are the selective binding to $R_1^{med}R_2^{hi}$ over $R_1^{hi}R_2^{med}$, and $R_1^{hi}R_2^{lo}$ over $R_1^{med}R_2^{lo}$ (Fig. 2a). For these two pairs, the benefit of affinity changes is limited by the relative discrepancy in receptor abundances (Fig. 2c,d). Since the ratios of receptor expressions between $R_1^{med}R_2^{hi}$ and $R_1^{hi}R_2^{med}$ are significantly lower than those between $R_1^{hi}R_2^{lo}$ and $R_1^{lo}R_2^{hi}$, the greatest binding selectivity that can be achieved for these two populations is also significantly lower. When both receptors are uniformally more abundant in a target population than the off target population, such as when comparing $R_1^{hi}R_2^{hi}$ to $R_1^{med}R_2^{med}$, affinity tuning fails to modulate binding selectivity (Fig. 2d). Therefore, to use affinity changes for selectivity enhancement, it is critical to identify which receptors a cell population of interest expresses the most uniquely as compared to off-target populations.

Our binding contour plots are also useful for considering the interplay between affinity modulation and intrapopulation receptor expression heterogeneity. For example, $R_1^{med}R_2^{med}$ has relatively high variance in both R_1 and R_2 expression (Fig. 2a). When the binding affinities to R_1 and R_2 are divergent, as shown in the subplots in the upper left corner and bottom right corner of Fig. 2a, the ellipse representing its range of receptor abundances rides through multiple contour lines, indicating wide variation in the quantity of bound ligand. This intrapopulation variation in the amount of ligand bound, however, is not observed when the affinities to R_1 and R_2 are more balanced. Thus, a population's receptor expression heterogeneity and variation are important factors to consider when modulating ligand affinity.

Valency Enables Selectivity Based on Quantitative Differences in Receptor Abundance

Given the limitations of deriving selectivity from affinity changes, we next explored the effect of valency

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC BY 4.0 International license.

changes (Fig. 3). Our previous work has shown that this binding model can accurately predict the multivalent binding between IgG antibodies and Fcγ receptors.[17] To further demonstrate our model's capacity to predict multivalent binding activity, we fit our model to a set of published experimental measurements wherein fluorescently labeled nanorings were assembled with specific numbers of binding units.[18] After fitting to determine the crosslinking coefficient of multivalent nanorings (K_x^*) and receptor affinity values (K_a), our model was able to accurately match the binding of nanorings with four or eight subunits to cells expressing a known abundance of receptor partners (Fig. S1a).

Multivalent ligand binding differs from monovalent binding in its nonlinear relationship with receptor density, allowing multivalent ligands to potentially selectively target cells based on receptor abundance.[19] The effects of varying ligand valency are visualized in Fig. 3a. Here, the ligand only binds to receptor R_1, not receptor R_2, and the valency of the ligand is varied across subplots. Varying ligand valency results in a nonlinear relationship between receptor abundance and binding magnitude. For example, in the monovalent case, there are roughly the same amount of contour lines between $R_1^{lo}R_2^{lo}$ and $R_1^{med}R_2^{lo}$ as there are between $R_1^{med}R_2^{lo}$ and $R_1^{hi}R_2^{lo}$. However, in the tetravalent case, there are comparatively more contour lines between $R_1^{med}R_2^{lo}$ and $R_1^{hi}R_2^{lo}$, indicating that ligand bound is a nonlinear function of receptor expression when considering multivalent binding (Fig. 3a).

Selectivity derived by changing valency requires coordinate changes in affinity. As many previous studies have suggested, multi-valent ligands can demonstrate particular selectivity to target populations with high receptor abundances when their subunits have low affinity to the receptors.[13,18] Comparing binding between cell populations exposed to ligands of variable affinities over a range of ligand valency demonstrates the interplay between these factors (Fig. 3b-d). Here, the ligand binding ratios between cell populations are shown for ligands of high, medium, and low affinities for R_1 (K_d of 1000, 100, and $10\ \mathrm{nM}$), and the shaded areas indicate the variance caused by intrapopulation heterogeneity, determined by sigma point filters. The ligand binding ratio between $R_1^{hi}R_2^{lo}$ and $R_1^{med}R_2^{lo}$ is maximized by low affinity ligands, but requires greater valency to achieve peak binding selectivity when compared to ligands of greater affinity (Fig. 3b). A similar valency optimum for a given affinity is seen for binding selectivity between $R_1^{hi}R_2^{hi}$ and $R_1^{med}R_2^{med}$, and $R_1^{hi}R_2^{med}$ and $R_1^{med}R_2^{med}$ (Fig. 3c-d). Ligands with lower affinities achieve optimal binding with higher valencies and exhibit higher selectivity for cells expressing greater amounts of receptor.

For any specific multivalent ligand at equilibrium, it may bind to any number of receptors up to its valency, which we will refer to as its binding degree. Our model allowed us to identify the mechanism of valency-mediated receptor abundance selectivity by examining the distribution of ligands bound at each degree to different cells for octavalent ligands (valency = 8). A cell expressing 10^4 receptors displays similar amounts of binding at each degree for ligands with dissociation constants of 1000, 100, and $10\ \mathrm{nM}$ (Fig. 3e). However, a cell expressing 10% as many receptors exhibits extremely low amounts of higher-degree binding for ligand complexes of low binding affinity (Fig. 3f). This effect arises due to the ligand's rate of dissociation being larger than its multivalent binding association rate at low receptor abundances (Fig. 3g). Multivalent ligands undergo initial binding events at rates unaffected by steric effects and changes in local concentration, but low affinity and receptor density severely limit the possibility of secondary binding events. In contrast, cells with a higher abundance of receptors accumulate ligand bound by having them bind multivalently, as the forward rate of

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

secondary binding events is greater than that of receptor-ligand dissociation (Fig. 3f). This effect allows multivalent ligands to achieve selective binding to cells based on their receptor abundances, which monovalent ligands, even with engineered affinities, are unable to do.

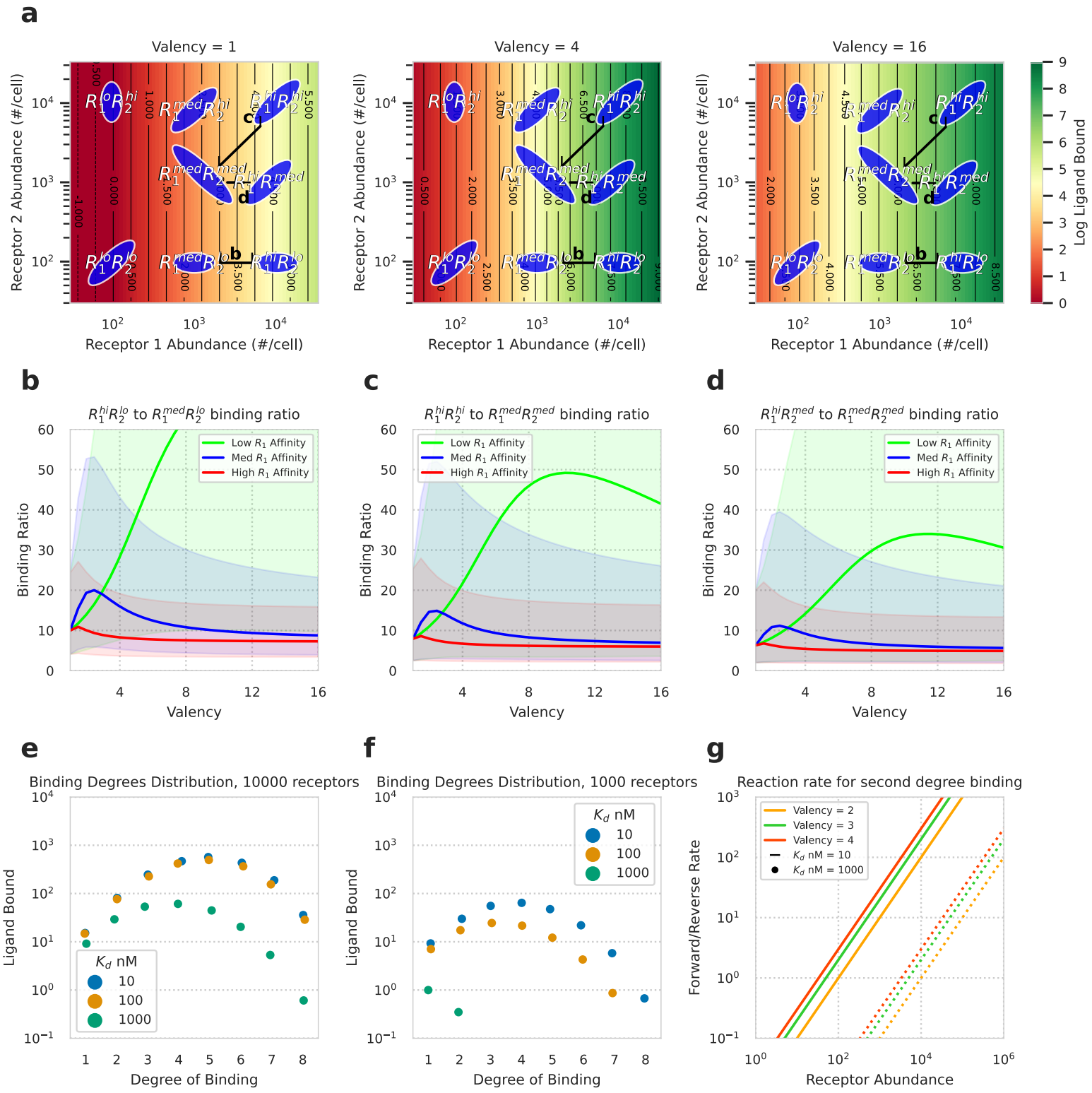


Figure 3: Valency provides selectivity based on receptor expression levels. a) Heat/contour maps of multivalent ligands bound to cell populations given their expression profiles of two receptors. Multivalent ligand subunits bind to only R_1 with an dissociation constant of $100\,\mathrm{nM}$, and do not bind to R_2. Complexes vary in valency from 1 to 16. Ligand concentration $L_0 = 1\,\mathrm{nM}$; crosslinking constant $K_x^* = 10^{-10}$. b-d) Ligand binding ratio between various cell populations for ligands of valency ranging from 1 to 16. The shaded areas indicate the variance of binding ratios caused by the intrapopulation heterogeneity and estimated by sigma point filters. b) Ligand bound ratio of $R_1^{hi}R_2^{lo}$ to $R_1^{med}R_2^{lo}$, c) $R_1^{hi}R_2^{hi}$ to $R_1^{med}R_2^{med}$, and

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-4.0 International license.

d) $R_1^{hi}R_2^{med}$ to $R_1^{med}R_2^{med}$. e-f) Number of ligand bound to each possible number of receptors for cells exposed to octavalent ligand complexes composed of subunits with dissociation constants of 1000, 100, or $10\,\mathrm{nM}$ for receptor 1. e) Number of octavalent complex bound at each degree for a cell with 10^4 receptors. f) Number of octavalent complex bound at each degree for a cell with 10^3 receptors. g) Ratio of forward to reverse binding rate for secondary binding events for multivalent ligands to cells expressing variable amounts of receptors.

Non-Overlapping Ligand Targeting Drives a Limited Selectivity Enhancement of Mixtures

While most therapies rely on the action of a single molecular species, mixtures may enhance selectivity through combinations of actions.[20] Furthermore, some biologics inevitably act as mixtures of species through heterogeneity in glycosylation and the presence of endogenous ligands.[21] Therefore, it is important to understand how mixtures of complexes influence overall response.

To evaluate the contribution of mixtures, we evaluated model-predicted binding while varying the composition between two distinct monovalent ligands, each exhibiting preferential binding to either R_1 or R_2(Fig. 4a). The trend that arises is very similar to an additive combination of the single ligand cases. This pattern highlights a key limitation of using mixtures for selectivity: selectivity between two populations varies monotonically with the composition, so any mixture combination is no better than using the more specific ligand entirely (Fig. 4b).

While mixture engineering fails to enhance binding selectivity between two cell populations, it is potentially beneficial when considering two or more off-target cell populations. More specifically, when a target population expresses two target receptors, but off-target populations express each receptor individually in high amounts, drug mixtures can offer enhanced selectivity. For example, when maximizing targeting to $R_1^{hi}R_2^{hi}$ over $R_1^{hi}R_2^{lo}$ and $R_1^{lo}R_2^{hi}$, which individually express high levels of the receptors found on the $R_1^{hi}R_2^{hi}$, a uniform mixture of ligands with high affinity for receptor 1 and 2 provides a modest improvement in targeting selectivity (Fig. 4c). However, even in these cases, the magnitude of selectivity enhancement is modest. Finally, although we only consider the amount of binding, ligands can have non-overlapping signaling effects even with identical amounts of binding. In these cases, the effect of combinations can be distinct from either individual ligand.[17,22]

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

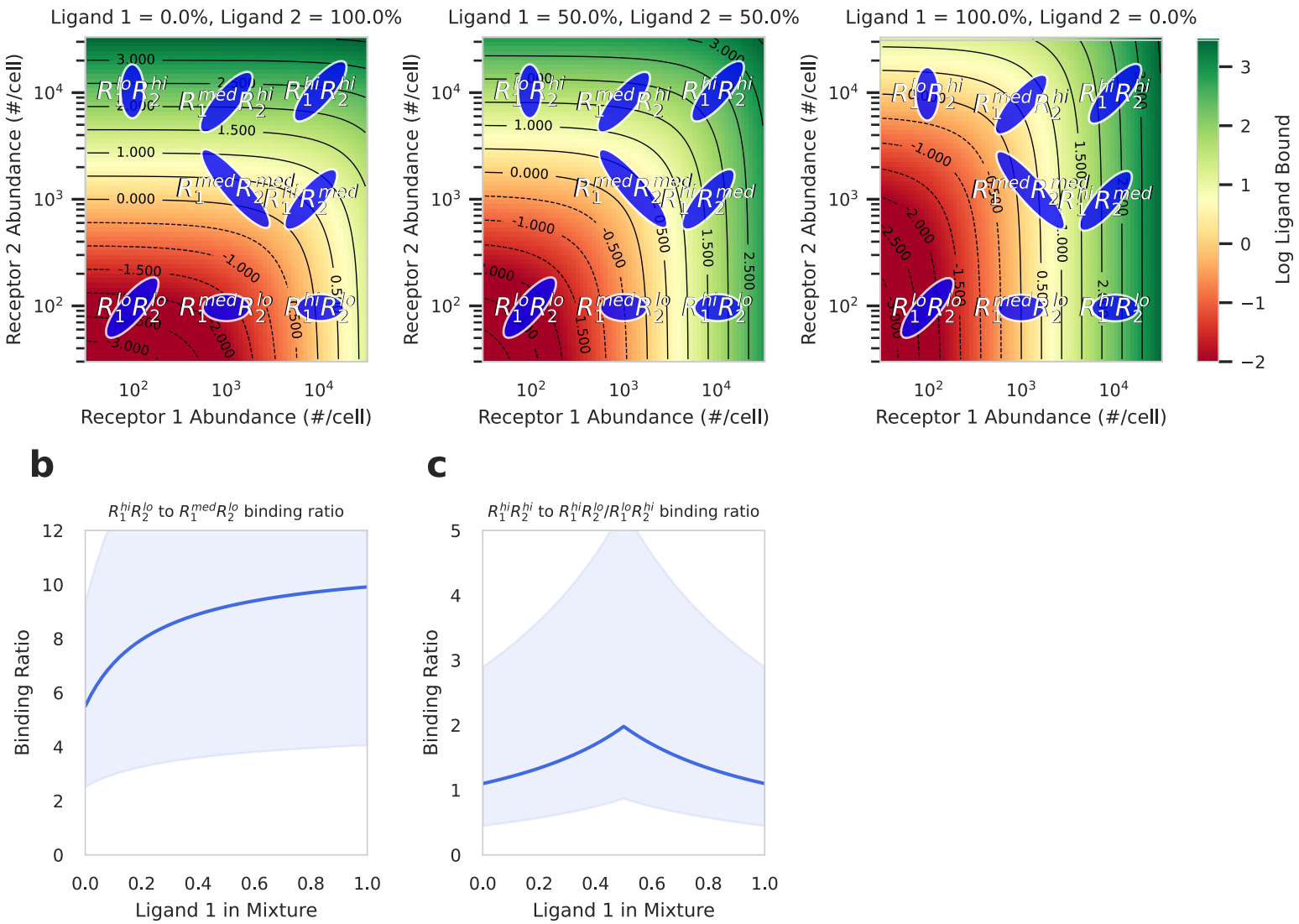


Figure 4: Ligand mixtures with non-overlapping responses can enhance selectivity. a) Heat/contour maps of multivalent ligands bound to cell populations given their expression profiles of two receptors. A mixture of monovalent ligands is used, with ligand 1 binding to receptor 1 and 2 with dissociation constants of $1\,\mu\mathrm{M}$ and $10\,\mu\mathrm{M}$ respectively, and ligand 2 binding to receptors 1 and 2 with dissociation constants of $10\,\mu\mathrm{M}$ and $1\,\mu\mathrm{M}$ respectively. Ligand concentration $L_0 = 1\,\mathrm{nM}$; crosslinking constant $K_x^* = 10^{-10}$. b,c) Ratio of ligand bound to cell populations exposed to monovalent mixtures of ligand 1 and 2. The ratio of the target population to the single off target population with the greatest ligand bound is plotted. The shaded areas indicate the variance caused by intrapopulation heterogeneity, estimated by sigma point filters. b) Ligand bound ratio of $R_1^{hi}R_2^{lo}$ to $R_1^{med}R_2^{lo}$, c) $R_1^{hi}R_2^{hi}$ to $R_1^{hi}R_2^{lo}$ and $R_1^{lo}R_2^{hi}$.

Heterovalent Bispecific Ligands Exhibit Unique Charateristics When Activated Fully Bound

Constructing multispecific drugs has become a promising new strategy for finer target cell specificity with the advancement of engineering techniques.[23] However, the number of possible configurations of multispecific drugs is combinatorially large and impossible to enumerate. Here, we use heterovalent bispecific ligands as examples to explore the unique benefit of multispecificity distinct from any strategy analyzed before. Here, we compare bispecificity with a 50%-50% mixture of two monovalent ligands, and a 50%-50% mixture of two different homogeneous bivalent ligands (Fig. 5i). These two strategies both have some similarities to bispecific therapeutics. A bispecific ligand contains two different ligand monomers with equal proportion, similar to a 50%-50% monovalent mixture. By

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

comparing bispecific with a 50%-50% monovalent mixture, we can elucidate the extra benefit of tethering these two monomers into one complex. A bispecific ligand is also by nature bivalent, and by comparing it with homogeneous bivalent drugs, we can understand how having two different subunits in the same complex modify the behavior of a drug. Together, we sought to identify the unique advantages afforded by heterovalent ligands.

We first applied the binding model to predict the amount of ligand bound in bispecific drugs (Fig. 5a), a 50%-50% mixture of two monomers (Fig. 5b), and a 50%-50% mixture of two different homogeneous bivalents (Fig. 5c), with the same set of parameters in ligand concentration and affinities. Surprisingly, the patterns of ligand binding in these three cases are almost exactly the same, and bispecificity appeared to offer no unique properties. However, many bispecifics only impart their therapeutic action when both of their subunits bind to a target population. For example, in the design of bispecific antibodies, it is common to require both subunits bound to be effective.[24,25] When the first antigen-binding region of a bispecific antibody docks at an antigen, it will only make a loose connection between the target and the antibody, unable to induce a strong immune response, and only when both antigen-binding regions bind to their target will this bispecific antibody be fully effective. We investigated whether bispecific ligands that have to be activated with both of their subunits binding to a target display any special cell population selectivity characteristics.

To scrutinize the case for bispecific ligands where the target binding of both subunits is required, we extended our model to calculate the amount of ligand fully bound (Fig. 5i). For the heterovalent bispecific case, ligand fully bound will only account for the amount of ligand with both of their ligand monomers bound to a target. With the same set of parameters, the predictions made for bispecific fully bound show a very distinct pattern from general ligand bound (Fig. 5e). The contour plot of fully bound bispecific ligands has more convex contour lines. For example, $R_1^{hi}R_2^{hi}$ has about the same level of general ligand bound as $R_1^{hi}R_2^{med}$ (Fig. 5a), but it has significantly more ligand fully bound than $R_1^{hi}R_2^{med}$ (Fig. 5e). This convexity of contour lines indicates that for bispecific complexes, double-positive cells bind more ligands fully than cells only strongly expressing one type of receptors. This should be obvious since these two subunits of the bispecific complex prefer different receptors. The results indicate that bispecific ligands will only exhibit special characteristics when it requires both receptor binding subunits to bind a receptor.

The specific amount of fully bound ligands is dependent on the ligand's propensity for crosslinking captured by the constant K_x^* (Fig. 5i). Typically in biochemistry, when there is less steric hindrance among the subunits of a multivalent drug molecule (e.g. longer tether and smaller subunit size), or when there is local receptor clustering on the target cell, secondary binding will be easier to achieve, corresponding to larger K_x^*.[26] We plotted the pattern of bispecific fully bound with $K_x^* = 10^{-10}$, 10^{-12}, and 10^{-14} (Fig.5d-f). In general, when K_x^* is larger, the ligands are more capable of multimerization, and there will be more fully bound units. To demonstrate how this characteristic of fully bound bispecific ligand imparts cell population specificity, we compared the selectivities between some chosen target-to-off-target cell population pairs under bispecific drug versus another drug given a range of K_x^* (Fig. 5g,h). These plotted numbers are the selectivities imparted by a bispecific drug divided by the selectivities from a drug mixture of either monovalent (Fig. 5g) or homogeneous bivalent (Fig. 5h). When these quotients are larger, it implies that fully bound bispecific ligand has greater advantages than its counterpart to impart selective binding for the chosen target cell

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

populations. Figure 5g compares the selectivities under bispecific ligands versus a 50%-50% mixture of monovalent ligands. The results show that fully bound bispecific can grant better selective binding when K_x^* is small enough. This is logical, since when K_x^* is small and crosslinking is rarer, most ligands will bind monovalently, and fully bound bispecific will especially favor the cell populations with higher receptor expression. However, when we compare fully bound bispecific to fully bound homogeneous bivalent mixtures (Fig. 5h), the advantage of bispecific drugs is not obvious except for $R_1^{med}R_2^{hi}$ to $R_1^{hi}R_2^{med}$ selectivity. Given that we only account for fully bound ligand for both therapeutics, the effect demonstrated in Figure 5g no longer holds. Together, we showed that bispecific ligands only exhibit unique advantages in inducing selective binding when they are only effective when both of their subunits bind and crosslinking is more difficult.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

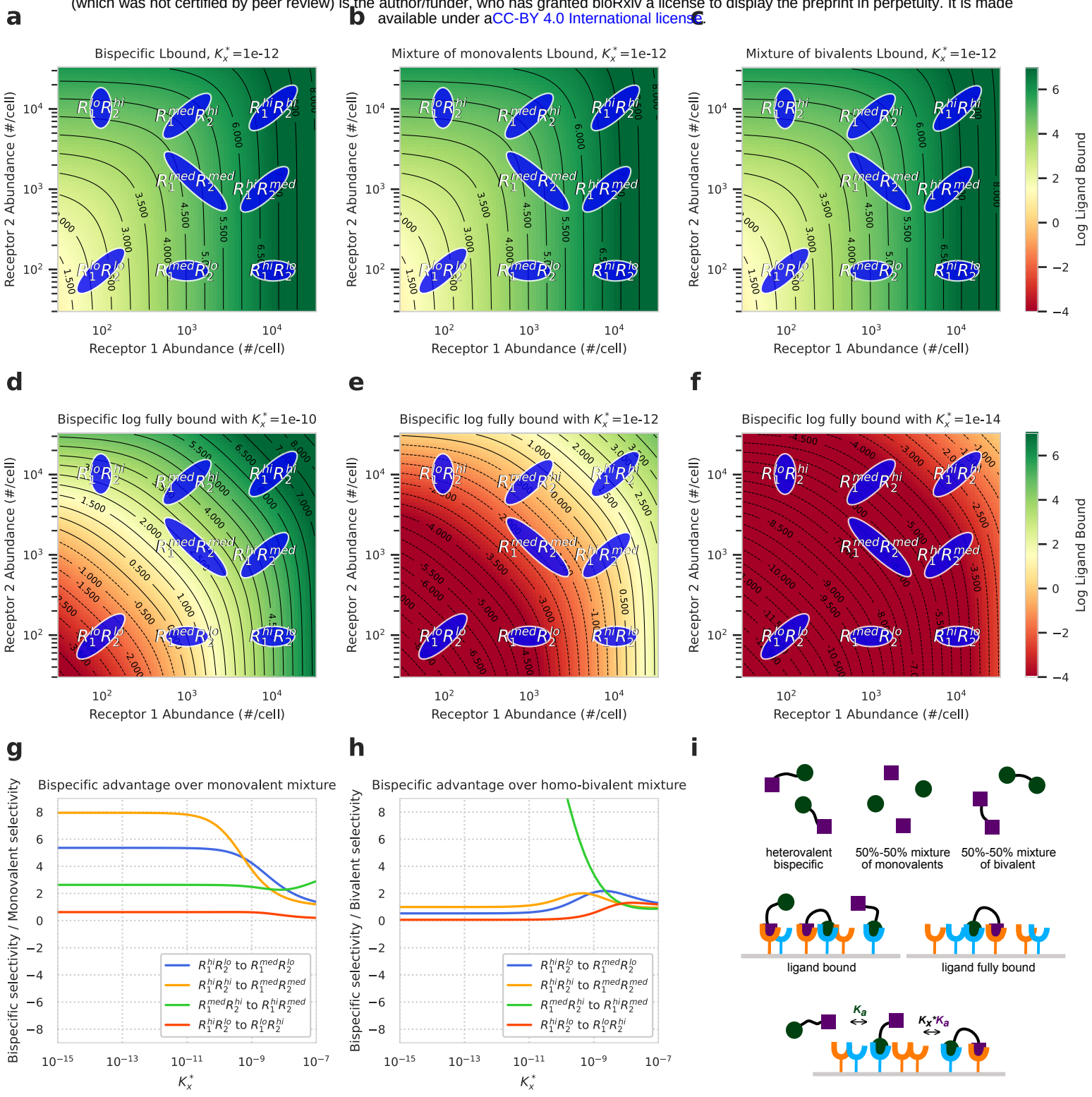


Figure 5: Bispecific ligands exhibit unique effects when they can only be activated with both subunits bind. a-c) Ordinary ligand bound are dominated by the initial binding of ligands, so it doesn't provide unique advantages to a) bispecific ligand comparing with b) a 50%-50% mixture of monovalent ligands and c) a 50%-50% mixture of bivalent ligands. Ligand concentration $L_0 = 10\text{nM}$; binding affinities $K_{d11} = 100\text{nM}$, $K_{d22} = 1\mu\text{M}$, $K_{d12} = K_{d21} = 10\mu\text{M}$. d-f) The amount of fully bound bispecific ligands depends on the tendency of multimerization, capsuled by K_x^*. d) $K_x^* = 10^{-10}$, e) $K_x^* = 10^{-12}$, f) $K_x^* = 10^{-14}$. g,h) Comparing bispecific selectivities with mixture selectivities, varies with K_x^*, the crosslinking constant. When the ratios are larger, bispecific ligands bind to target populations more specifically. g) bispecific selectivities divided by monovalent 50%-50% mixture selectivities, h) bispecific selectivities divided by a 50%-50% homo-bivalent mixture selectivities. i) Cartoons of bispecific ligands and ligand fully bound binding model.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Combining Strategies For Superior Selectivity

Each strategy described above provided selectivity benefits in distinct situations, suggesting that they might synergistically improve selectivity when combined. We explored this potential synergy using optimization to determine the ligand specifications which provided optimal selectivity for our theoretical cell populations. More specifically, we optimized the selectivity of a ligand for a particular population, while considering all other populations to be off-target. Our optimization allowed ligand characteristics to vary within biologically plausible bounds; which characteristics we allowed to vary defined our examination of the efficacy of our various ligand engineering strategies. We examined optimizing affinity alone, mixture along with affinity, and valency along with affinity, and finally combined all three strategies (Fig. 6).

Optimizing a ligand for selectivity to $R_1^{lo}R_2^{hi}$ highlights a situation in which affinity imparts greater specificity, and optimal selectivity is achieved by combining affinity and valency modulations (Fig. 6a-f). Here selectivity is optimized by ligands with selective binding to receptor 2, and those with valencies allowing for selective binding to cells with higher abundances of receptor expression. One case contradictory to this trend is shown during the optimization for selectivity towards the $R_1^{hi}R_2^{hi}$ (Fig. 6g-l). While affinity engineering is unable to impart some small contributions to enhanced selectivity, significant improvement is only achieved when utilizing valency modulation techniques. A more difficult design problem is featured in the optimization of $R_1^{med}R_2^{med}$ (Fig. 4m-r). Since it lies in the midst of the other populations in receptor expression space, any modulation of affinity, valency or combining it with mixture-based strategies seems ineffective. It should be noted that in all described cases, engineering the composition of a mixture of ligands is generaly ineffective for imparting selectivity when the ligand's design specifications are flexible, and is likely only efficacious when using ligands with static properties and considering multiple off-target populations.

Our results highlight that both in singular and combined strategies for therapeutic manipulation, the target and off-target populations dictate the optimal approach. It is also clear that combined approaches do offer synergies which can be harnessed, but that those are only emergent in particular therapeutic situations.[25]

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

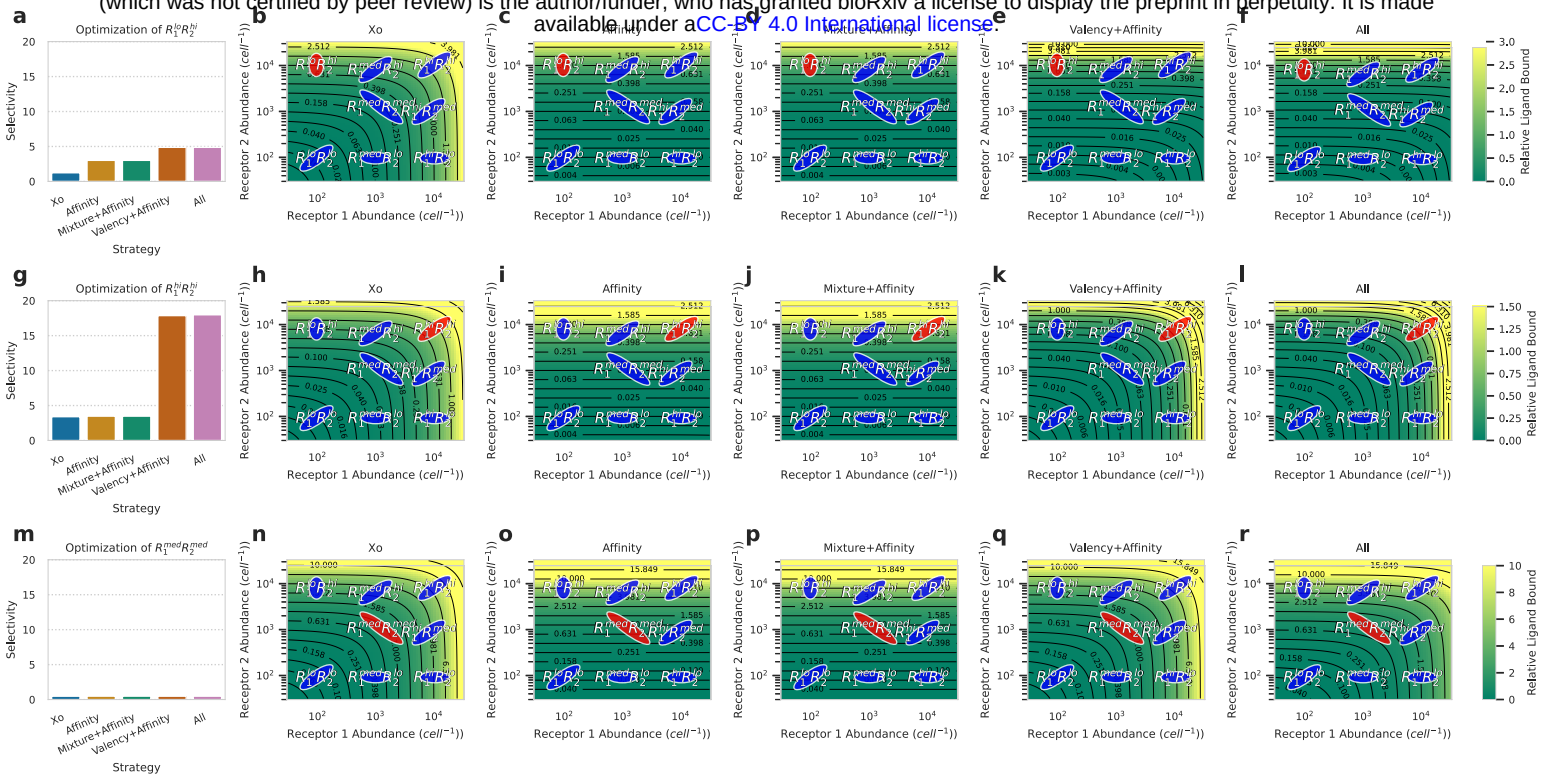


Figure 6: Combinations of strategies provide superior selectivity. a,g,m) Optimal selectivity levels (ligand bound on target population divided by average ligand bound by all other populations) achieved using various ligand engineering techniques. Ligand concentration $L_0 = 1\text{nM}$. Xo ligands are monovalent ligands with affinities of $1\ \mu\text{M}$ for both receptor 1 and 2. The dissociation constant was allowed to vary between $10\ \text{mM}$ and $0.1\ \text{nM}$ for both receptors using the "affinity" approach. Valency was allowed to vary from 1 to 16 for the "valency" approach in addition to affinities varying. Mixtures were assumed to be composed of two monovalent ligands, and affinities were allowed to vary in the "mixture" approach. The combined "all" approach allowed all of these quantities to vary simultaneously. The crosslinking constant K_x^* was allowed to vary between 10^{-15} and 10^{-9} for all approaches. b-f,h-l,n-f) Heatmap of magnitude of ligand bound for ligand with optimized characteristics according to various ligand engineering strategies. Target population is shown in red. a-f) Pertains to optimal targeting of $R_1^{lo}R_2^{hi}$, g-l) pertains to optimal targeting of $R_1^{hi}R_2^{hi}$, and m-r) pertains to optimal targeting of $R_1^{med}R_2^{med}$.

Using Binding Competition to Invert Receptor Targeting

While the strategies above provided selectivity in many cases, we recognized that they are all limited to a positive relationship between receptor abundance and binding. Therefore, we wondered if binding competition with a receptor antagonist, or "dead ligand", could invert this relationship.

To investigate the effect of ligand competition with an antagonist, we modeled mixtures of ligands, but only quantified the amount of binding for the active ligand. Here we chose to only consider a monovalent agonist and tetravalent antagonist (Fig. 7). We found that combinations of monovalent agonistic ligands and multivalent antagonistic ligands were able to uniquely target populations expressing small or intermediate amounts of receptors, which is demonstrated when comparing ligand binding ratios between $R_1^{med}R_2^{lo}$ to $R_1^{hi}R_2^{lo}$ (Fig. 7a). Here, a nearly tenfold increase in selectivity can be granted to monovalent agonists when combined with a tetravalent antagonist. In this case, there are greater quantities of agonist bound to $R_1^{med}R_2^{lo}$ than $R_1^{hi}R_2^{lo}$ (Fig. 7b). This is striking, as $R_1^{med}R_2^{lo}$ expresses either as many or fewer abundances of receptors one and two when compared to $R_1^{hi}R_2^{lo}$. This phenomenon, which could not be achieved without multivalent antagonists, occurs due to the preferential binding of multivalent antagonists to populations expressing higher abundances of

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

receptors (Fig. 7c, 3e). Thus, in cases where previously discussed ligand engineering strategies and approaches fail to achieve selective binding to cells expressing smaller or similar amounts of receptors to off-target populations, combinations of agonistic and antagonistic ligands may provide unique benefits.

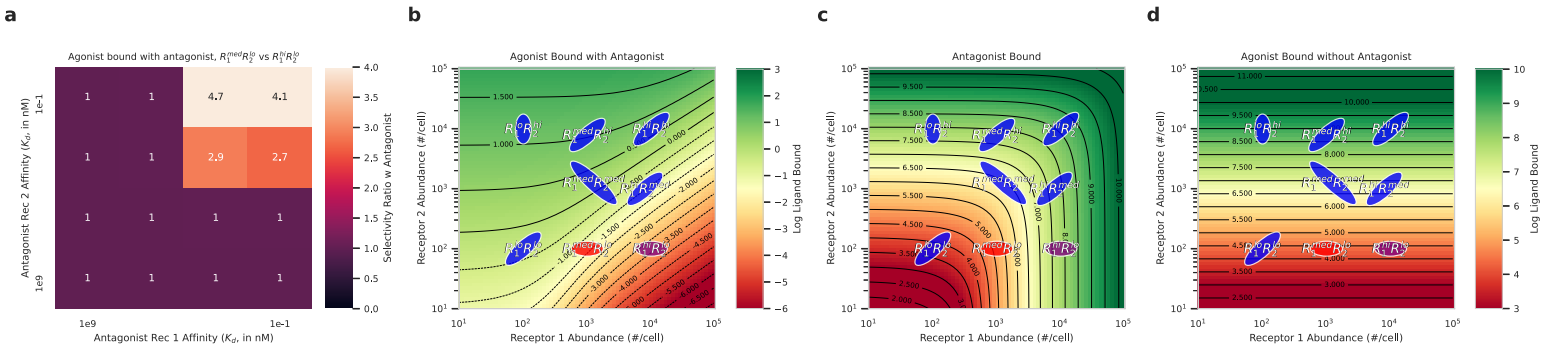


Figure 7: Mixtures of receptor agonists and antagonists allow for unique population targeting activity. Ligand concentration $L_0 = 1\text{nM}$. a) Selectivity for $R_1^{med}R_2^{lo}$ against $R_1^{hi}R_2^{lo}$ when exposed to a tetravalent "dead ligand" antagonist with varying affinities for receptors 1 and 2, and a monovalent therapeutic receptor agonists with affinities optimized for selectivity. Only amount of agonist bound is considered in determination of optimal selectivity. b-d) Heatmap of agonist (b), and antagonist (c) ligand bound for antagonist and agonist ligand combination shown to impart greatest selectivity improvement in (a). d) Heatmap of agonist bound in b,c when no antagonist is present.

Discussion

Here, we developed and employed a multivalent, multi-ligand, multi-receptor binding model and used it to explore the effectiveness of various ligand engineering strategies for population-selective binding (Fig. 1). Using a representative set of theoretical cell populations defined by their distinct expression of two receptors, we examined the efficacy of several potential ligand engineering strategies, including changes to affinity, ligand valency, mixtures of species, multi-specificity, antagonist competition, and these in combination.

Each strategy's contribution can be summarized by general patterns. We found that affinity changes were most effective when the target and off-target populations expressed distinct combinations of receptors (Fig. 2). Binding selectivity was enhanced by increasing the affinity of the ligand for those receptors that are more abundantly expressed on the target cells and decreasing its affinity for those that are not. Selectivity depended upon large differences in receptor expression levels between on- and off-target populations. When target and off-target populations expressed the same pattern of receptors and only differed in receptor abundances, modulations in valency, but not affinity, were effective (Fig. 3). A key determinant of valency's effectiveness was the secondary binding and unbinding rate which is dependent on both the kinetics of the receptor-ligand interaction and receptor abundance. Ligand mixtures were mostly ineffective for imparting binding selectivity, and only had modest benefits when considering two or more off-target populations (Fig. 4). Heterovalent bispecific ligands only showed unique advantages over mixtures of monovalent ligands or bivalent ligands when solely considering fully bound ligand (Fig. 5). These ligands exhibit preferential binding to target populations that have high expression of both receptors over those with high expression of only a single receptor, with the propensity for secondary binding acting as a key determinant for selectivity. We found that, while a single ligand engineering strategy dominated in its contributions to cell type

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

selectivity, synergies between these strategies existed in some cases to derive even greater specificity (Fig. 6). Finally, we found that combinations of monovalent therapeutic ligands with multivalent antagonistic ligands allow for the unique selection of target populations expressing relatively fewer receptors than off-target populations (Fig. 7).

While our multivalent binding model identified strategies for selective targeting in many cases, it also identified situations for which selective binding is challenging. For example, selectively targeting populations based on their absence of receptor expression remains challenging. While we computationally show the potential of using multivalent antagonists with monovalent agonists to selectively target such populations, implementing this may be challenging in practice. In cases where a target population expresses fewer receptors of any kind than an off-target population, our analysis suggests that targeting other receptors should be considered. However, in cases where target populations express more of any type of receptor than an off-target population, we show that one or more of our formulated ligand engineering strategies can be employed to improve binding selectivity. While we expect the same patterns to apply with greater than two receptors, certain emergent behaviors may exist with tri-specific and more complex ligand binding.

A few of the strategies that we explored have been utilized in existing engineered therapies. For example, affinity changes to the cytokine IL-2 have been used to bias its effects towards either effector or regulatory immune populations.[27,28] Varying the valency of tumor-targeted antibodies leads to selective cell clearance based upon the levels of expressed antigen.[19] Manipulating of the affinities of the fibronectin domains on octovalent nanorings was shown to enhance the selectivity of binding to cancerous cells displaying relatively higher densities of fibronectin receptors compared to native tissue.[18] The tendency of low-affinity, multivalent interactions to target cells expressing high receptor abundances was also described in a study describing the selectivity of multivalent antibody binding to tumor cells bound by a bispecific therapeutic ligand.[13] These examples lend support to the accuracy and translational value of our model. At the same time, recognizing these previously described ligand engineering approaches as separable strategies provides clearer guidance for future engineering.

Some of the optimization strategies described here have not been exploited in part due to the complexity of real biological applications. First, some strategies may not be biochemically practical. For example, the manipulating ligand affinity requires intricate protein engineering. Potential dynamic changes in the receptor expression profile of a target population also complicate the matter. It is well documented that cancer cells can escape therapeutic targeting by upregulating[29,30] or downregulating[31] the expression of certain receptors. In this case, both the current and potential abundance of each receptor must be considered. While this work does not address many of these issues, we propose that using a computational binding model can analyze these strategies quantitatively and collectively from a mechanistic perspective. Even when the absolute mathematically optimal ligand characteristics cannot be achieved biochemically, our analyses provide guidance within the bounds of what is attainable and how to approach the optimum, accounting for implementation feasibility and facilitating the implementation of strategy combinations.

In many therapeutic applications where selective engagement of target cell populations is an important performance metric, such as the treatment of cancer, cellular therapies are becoming increasingly popular.[32] Human engineered chimeric antigen receptor (CAR) T cells have enhanced the potential to

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

selectively recognize and attack malignant tissues.[14] These technologies bypass ligand-receptor binding restrictions by allowing recognition in signaling response. However, we have shown here that selectivity can often be attained with relatively simple therapeutic ligands. This study lays the framework for how ligand engineering can be directed using computational modeling. It should be noted that the application of this logic is reliant on knowledge of the target and off-target cell population receptor expression levels. Future application of the ligand binding logic described in this study could be guided using high-throughput single-cell profiling techniques, such as RNA-seq or high-parameter flow cytometry. A computational tool that could directly translate such datasets into ligand design criteria, selecting among potential receptor targets, may represent a potential avenue for the translation of our analyses into a more broadly applicable ligand engineering tool.

Materials and Methods

Data and Software Availability

All analysis was implemented in Python v3.8, and can be found at https://github.com/meyer-lab/cell-selective-ligands.

Generalized multi-ligand, multi-receptor multivalent binding model

To model multivalent ligand complexes, we extended our previous binding model to the multi-ligand case.[17] We define N_L as the number of distinct monomer ligands, N_R the number of distinct receptors, and the association constant of monovalent binding between ligand i and receptor j as $K_{a,ij}$. Multivalent binding interactions after the initial interaction have an association constant of $K_x^* K_{a,ij}$, proportional to their corresponding monovalent affinity. The concentration of complexes is L_0, and the complexes consist of random ligand monomer assortments according to their relative proportion. The number of ligand complexes in the solution is usually much greater than that of the receptors, so we assume binding does not deplete the ligand concentration. The proportion of ligand i in all monomers is C_i. By this setup, we know $\sum_{i=1}^{N_L} C_i = 1$. $R_{\text{tot},i}$ is the total number of receptor i expressed on the cell surface, and $R_{\text{eq},i}$ the number of unbound receptors i on a cell at the equilibrium state during the ligand complex-receptor interaction.

The binding configuration at the equilibrium state between an individual complex and a cell expressing various receptors can be described as a vector $\mathbf{q} = (q_{1,0}, q_{1,1}, \ldots, q_{1,N_R}, q_{2,0}, \ldots, q_{2,N_R}, q_{3,0}, \ldots, q_{N_L,N_R})$ of length $N_L(N_R + 1)$, where $q_{i,j}$ is the number of ligand i bound to receptor j, and $q_{i,0}$ is the number of unbound ligand i on that complex in this configuration. The sum of elements in $\mathbf{q}$ is equal to f , the effective avidity. For all i in $\{1, 2, \ldots, N_L\}$, let $\varphi_{i,j} = R_{\text{eq},j} K_{a,ij} K_x^* C_i$ when j is in $\{1, 2, \ldots, N_R\}$, and $\varphi_{i,0} = C_i$. The relative amount of complexes in the configuration described by $\mathbf{q}$ at equilibrium is

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

$$v_{\mathbf{q},eq} = \binom{f}{\mathbf{q}} \frac{L_0}{K_x^*} \prod_{\substack{i=1 \\ j=0}}^{\substack{i=N_L \\ j=N_R}} {\varphi_{ij}}^{q_{ij}},$$

with $\binom{f}{\mathbf{q}}$ being the multinomial coefficient. Then the total relative amount of bound receptor type n at equilibrium is

$$R_{\text{bound},n} = \frac{L_0 f}{K_x^*} \sum_{m=0}^{N_L} \varphi_{mn} \left(\sum_{\substack{i=1 \\ j=0}}^{\substack{i=N_L \\ j=N_R}} {\varphi_{ij}}^{q_{ij}} \right)^{f-1}.$$

By conservation of mass, we know that $R_{\text{tot},n} = R_{\text{eq},n} + R_{\text{bound},n}$ for each receptor type n, while $R_{\text{bound},n}$ is a function of $R_{\text{eq},n}$. Therefore, each $R_{\text{eq},n}$ can be solved numerically using $R_{\text{tot},n}$. Similarly, the total relative amount of complexes bound to at least one receptor on the cell is

$$L_{\text{bound}} = \frac{L_0}{K_x^*} \left[\left(\sum_{\substack{i=1 \\ j=0}}^{\substack{i=N_L \\ j=N_R}} {\varphi_{ij}}^{q_{ij}} \right)^{f} - 1 \right].$$

Generalized multivalent binding model with defined complexes

When complexes are engineered and ligands are not randomly sorted into multivalent complexes, such as with the Fabs of bispecific antibodies, the proportions of each kind of complex become exogenous variables and are no longer decided by the monomer composition C_i's. The monomer composition of a ligand complex can be represented by a vector $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_{N_L})$, where each θ_i is the number of monomer ligand type i on that complex. Let $C_{\boldsymbol{\theta}}$ be the proportion of the $\boldsymbol{\theta}$ complexes in all ligand complexes, and Θ be the set of all possible $\boldsymbol{\theta}$'s. We have $\sum_{\boldsymbol{\theta} \in \Theta} C_{\boldsymbol{\theta}} = 1$.

The binding between a ligand complex and a cell expressing several types of receptors can still be represented by a series of q_{ij}. The relationship between q_{ij}'s and θ_i is given by $\theta_i = q_{i0} + q_{i1} + \ldots + q_{iN_R}$. Let the vector $\mathbf{q}_i = (q_{i0}, q_{i1}, \ldots, q_{iN_R})$, and the corresponding $\boldsymbol{\theta}$ of a binding configuration $\mathbf{q}$ be $\boldsymbol{\theta}(\mathbf{q})$. For all i in $\{1, 2, \ldots, N_L\}$, we define $\psi_{ij} = R_{\text{eq},j} K_{a,ij} K_x^*$ where $j = \{1, 2, \ldots, N_R\}$ and $\psi_{i0} = 1$. The relative amount of complexes bound to a cell with configuration $\mathbf{q}$ at equilibrium is

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-4.0 International license.

$$v_{\mathbf{q},eq} = \frac{L_0 C_{\boldsymbol{\theta}(\mathbf{q})}}{K_x^*} \prod_{\substack{i=1 \\ j=0}}^{\substack{j=N_R}} \psi_{ij}{}^{q_{ij}} \prod_{i=1}^{N_L} \binom{\theta_i}{\mathbf{q}_i}.$$

Then we can calculate the relative amount of bound receptor n as

$$R_{\text{bound},n} = \frac{L_0}{K_x^*} \sum_{\boldsymbol{\theta} \in \Theta} C_{\boldsymbol{\theta}} \left[\sum_{i=1}^{N_L} \frac{\psi_{in} \theta_i}{\sum_{j=0}^{N_R} \psi_{ij}} \right] \prod_{i=1}^{N_L} \left(\sum_{j=0}^{N_R} \psi_{ij} \right)^{\theta_i}.$$

By $R_{\text{tot},n} = R_{\text{eq},n} + R_{\text{bound},n}$, we can solve $R_{\text{eq},n}$ numerically for each type of receptor. The total relative amount of ligand binding at equilibrium is

$$L_{\text{bound}} = \frac{L_0}{K_x^*} \sum_{\boldsymbol{\theta} \in \Theta} C_{\boldsymbol{\theta}} \left[\prod_{i=1}^{N_L} \left(\sum_{j=0}^{N_R} \psi_{ij} \right)^{\theta_i} - 1 \right].$$

Mathematical optimization

We used the SciPy function `scipy.optimize.minimize` to combine several strategies and achieve optimal selectivity.[34] Unless specified otherwise, the initial values for optimization were 10^{-12} for crosslinking coefficient K_x^*, 1 for valency f, 100% ligand 1 for mixture composition, and $1\ \mu\text{M}$ for the affinity dissociation constants. The boundaries were 10^{-15} to 10^{-9} for K_x^*, 1–16 for f, 0–100% ligand 1 for mixture composition, and $10\ \text{mM}$ to $0.1\ \text{nM}$ for the affinity dissociation constants.

Sigma point filter

To consider the intrapopulation variance of a cell population in the optimization, we implemented the sigma point filters,[16] a computationally efficient method to approximate the variance proprogated through an ordinary differential equation-based model.

Reimplementation of Csizmar et al.

To validate our model, we recapitulated multivalent binding data from Csizmar et al. using our multivalent binding model (Fig. S1).[18] Here, fluorescently labeled nanorings displaying 1, 2, 4, and 8 fibronectin clones, which bind to epithelial cell adhesion molecule (EpCAM) antigens, were fabricated. Binding activity of nanorings displaying both high (C5) and low (B22) affinity EpCAM binding domains was measured. Binding to an EpCAM$^{\text{high}}$ (3.3×10^6 antigens/cell) population was measured using flow cytometry. We used nonlinear least squares optimization, as described above, to fit our multivalent binding model to the binding data, using a unique scaling factor for each fibronectin clone to convert between measured fluorescent intensity and magnitude of ligand bound. We allowed affinity of the fibronectin clones to vary during optimization.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Supplementary Figures

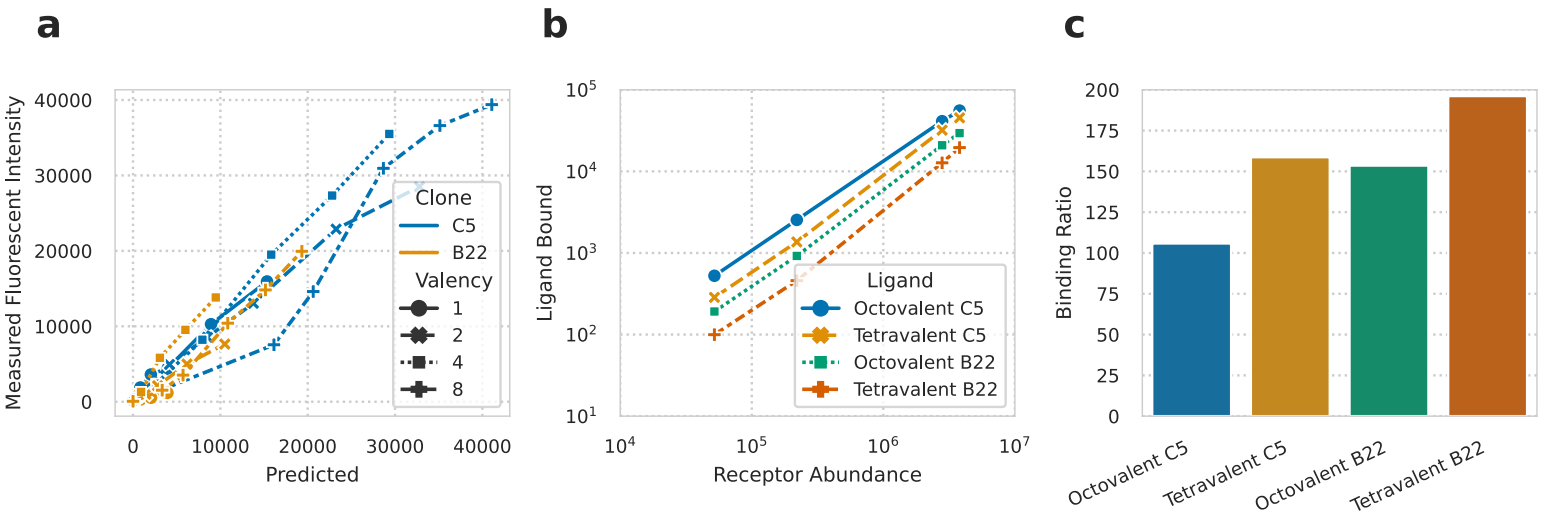


Figure 8: Our model is able to recapitulate experimental multivalent binding activity. a) Experimental vs. predicted fluorescent units for MCF-7 cells expressing 3.8×10^6 fibronectin receptors per cell bound to nanorings expressing one, two, four or eight fibronectin binding domains at concentrations of $0.16 – 500\,\mathrm{nM}$. C5 and B22 are high and low affinity fibronectin binding domains respectively. Using nonlinear least-squares regression, the crosslinking coefficient was found to be 1.11×10^{-12}; the dissociation constants for C5 and B22 were found to be $1.73\,\mu\mathrm{M}$ and $3.24\,\mu\mathrm{M}$ respectively, and ligand to fluorescent conversion factor for C5 and B22 were found to be 5.1×10^{-2} and 3.2×10^{-2} respectively. b) Predicted fluorescent values for MDA cells expressing 5.2×10^4, SK cells expressing 2.2×10^5, LNCaP cells expressing 2.8×10^6, and MCF-7 cells expressing 3.8×10^6 receptors per cell as described in Csizmar et al. [18]. Fluorescence was predicted for interactions with octa- and tetravalent ligands expressing either C5 or B22 fibronecting binding domains. c) Predicted binding ratios of MCF-7 cells to MDA cells when bound with octa- and tetravalent ligands expressing either C5 or B22 fibronecting binding domains.

Acknowledgements

This work was supported by NIH U01-AI148119 to A.S.M. **Competing financial interests:** The authors declare no competing financial interests.

Author contributions statement

A.S.M. concieved of the work. All authors implemented the analysis and wrote the paper.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

References

1. Mittendorf, E. A. *et al.* Loss of HER2 Amplification Following Trastuzumab-Based Neoadjuvant Systemic Therapy and Survival Outcomes. *Clinical Cancer Research* **15**, 7381–7388 (2009).

2. Turke, A. B. *et al.* Preexistence and clonal selection of MET amplification in EGFR mutant NSCLC. *Cancer Cell* **17**, 77–88 (2010).

3. Paek, A., Liu, J., Loewer, A., Forrester, W. & Lahav, G. Cell-to-Cell Variation in p53 Dynamics Leads to Fractional Killing. *Cell* **165**, 631–642 (2016).

4. Maude, S. L. *et al.* Chimeric Antigen Receptor T Cells for Sustained Remissions in Leukemia. *New England Journal of Medicine* **371**, 1507–1517 (2014).

5. O'Rourke, D. M. *et al.* A single dose of peripherally infused EGFRvIII-directed CAR T cells mediates antigen loss and induces adaptive resistance in patients with recurrent glioblastoma. *Science Translational Medicine* **9**, eaaa0984 (2017).

6. Shalek, A. K. *et al.* Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* **510**, 363–9 (2014).

7. Mitra, S. *et al.* Interleukin-2 activity can be fine tuned with engineered receptor signaling clamps. *Immunity* **42**, 826–38 (2015).

8. Chen, Z. *et al.* De novo design of protein logic gates. *Science* **368**, 78–84 (2020).

9. Srivastava, S. *et al.* Logic-Gated ROR1 Chimeric Antigen Receptor Expression Rescues T Cell-Mediated Toxicity to Normal Tissues and Enables Selective Tumor Targeting. *Cancer Cell* **35**, 489–503.e8 (2019).

10. Cho, J. H., Collins, J. J. & Wong, W. W. Universal Chimeric Antigen Receptors for Multiplexed and Logical Control of T Cell Responses. *Cell* **173**, 1426–1438.e11 (2018).

11. Ianevski, A., Giri, A. K. & Aittokallio, T. Fully-automated cell-type identification with specific markers extracted from single-cell transcriptomic data. *bioRxiv* (2019). doi:10.1101/812131

12. Robinson, M. K. *et al.* Targeting ErbB2 and ErbB3 with a bispecific single-chain Fv enhances targeting selectivity and induces a therapeutic effect in vitro. *British Journal of Cancer* **99**, 1415–1425 (2008).

13. Carlson, C. B., Mowery, P., Owen, R. M., Dykhuizen, E. C. & Kiessling, L. L. Selective Tumor Cell Targeting Using Low-Affinity, Multivalent Interactions. *ACS Chemical Biology* **2**, 119–127 (2007).

14. Silva, D.-A. *et al.* De novo design of potent and selective mimics of IL-2 and IL-15. *Nature* **565**, 186–191 (2019).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

15. Busse, D. *et al.* Competing feedback loops shape IL-2 signaling between helper and regulatory T lymphocytes in cellular microenvironments. *Proceedings of the National Academy of Sciences* **107**, 3058–3063 (2010).

16. Merwe, R. van der. Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models. (2004).

17. Robinett, R. A. *et al.* Dissecting FcγR Regulation through a Multivalent Binding Model. *Cell Syst* **7**, 41–48.e5 (2018).

18. Csizmar, C. M. *et al.* Multivalent Ligand Binding to Cell Membrane Antigens: Defining the Interplay of Affinity, Valency, and Expression Density. *Journal of the American Chemical Society* **141**, 251–261 (2018).

19. Mazor, Y. *et al.* Enhanced tumor-targeting selectivity by modulating bispecific antibody binding affinity and format valence. *Scientific Reports* **7**, (2017).

20. Lehár, J. *et al.* Synergistic drug combinations tend to improve therapeutically relevant selectivity. *Nature Biotechnology* **27**, 659–666 (2009).

21. Olivova, P., Chen, W., Chakraborty, A. B. & Gebler, J. C. Determination of N-glycosylation sites and site heterogeneity in a monoclonal antibody by electrospray quadrupole ion-mobility time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* **22**, 29–40 (2008).

22. Antebi, Y. E. *et al.* Combinatorial Signal Perception in the BMP Pathway. *Cell* **170**, 1184–1196.e24 (2017).

23. Deshaies, R. J. Multispecific drugs herald a new era of biopharmaceutical innovation. *Nature* **580**, 329–338 (2020).

24. Piccione, E. C. *et al.* A bispecific antibody targeting CD47 and CD20 selectively binds and eliminates dual antigen expressing lymphoma cells. *mAbs* **7**, 946–956 (2015).

25. Qi, J., Chen, S.-S., Chiorazzi, N. & Rader, C. An IgG1-like bispecific antibody targeting CD52 and CD20 for the treatment of B-cell malignancies. *Methods* **154**, 70–76 (2018).

26. Hlavacek, W. S., Posner, R. G. & Perelson, A. S. Steric effects on multivalent ligand-receptor binding: exclusion of ligand sites by bound cell surface receptors. *Biophys J* **76**, 3031–43 (1999).

27. Bentebibel, S.-E. *et al.* A First-in-Human Study and Biomarker Analysis of NKTR-214, a Novel IL2Rβγ-Biased Cytokine, in Patients with Advanced or Metastatic Solid Tumors. *Cancer Discovery* **9**, 711–721 (2019).

28. Peterson, L. B. *et al.* A long-lived IL-2 mutein that selectively activates and expands regulatory T cells as a therapy for autoimmune disease. *J Autoimmun* **95**, 1–14 (2018).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.398974; this version posted November 27, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

29. Towers, C. G. *et al.* Cancer Cells Upregulate NRF2 Signaling to Adapt to Autophagy Inhibition. *Developmental Cell* **50**, 690–703.e6 (2019).

30. Knowles, L. M. *et al.* Integrin v 3 and Fibronectin Upregulate Slug in Cancer Cells to Promote Clot Invasion and Metastasis. *Cancer Research* **73**, 6175–6184 (2013).

31. Bubeník, J. Tumour MHC class I downregulation and immunotherapy (Review). *Oncol Rep* **10**, 2005–8

32. Guedan, S., Ruella, M. & June, C. H. Emerging Cellular Therapies for Cancer. *Annual Review of Immunology* **37**, 145–171 (2019).

33. Rezvani, K., Rouce, R., Liu, E. & Shpall, E. Engineering Natural Killer Cells for Cancer Immunotherapy. *Molecular Therapy* **25**, 1769–1781 (2017).

34. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* (2020). doi:10.1038/s41592-019-0686-2