PNAS
www.pnas.org

Supplementary Information for:

Mechanisms of noncanonical binding dynamics in multivalent protein–protein interactions

Wesley J. Errington[a,1], Bence Bruncsics[b,1], and Casim A. Sarkar[a,2]

[a] Department of Biomedical Engineering, University of Minnesota, Minneapolis, MN 55455-0215

[b] Department of Measurement and Information Systems, Budapest University of Technology and Economics, H-1111 Budapest, Hungary

[1] These authors contributed equally to this work.

[2] To whom correspondence may be addressed: csarkar@umn.edu

This PDF file includes:

Other supplementary material for this manuscript includes the following:

Movie S1

Supplementary Text

Multivalent modeling: design and implementation

1. Conceptualization of a modeling framework

The objective of our model is to predict the binding kinetics of multivalent molecules across a broad region of parameter space. The conceptual framework consists of mathematically describing receptors and cognate ligands in a bead-on-a-string conformation with linkers spanning the individual receptor and ligand binding units. The model was formulated to generate interaction profiles mimicking the sensorgram outputs of surface plasmon resonance (SPR) experiments, a conventional and powerful method for quantifying receptor-ligand interaction kinetics (1).

2. Mass-action kinetic binding

The core of the binding model is the simple reversible ligand-receptor binding equilibrium, $R + L \rightleftharpoons RL$, with association and dissociation rate constants k_{on} and k_{off}, respectively. The relationships among the equilibrium concentrations, rate constants, and the equilibrium dissociation constant (K_D) are $K_D = \frac{k_{off}}{k_{on}} = \frac{[R][L]}{[RL]}$. For multivalent molecules, we assume that k_{on} and k_{off} represent intrinsic binding and dissociation propensities and are therefore the same for monovalent receptor-ligand interactions and for receptor and ligand units within a multivalent framework. To introduce a nomenclature for binding events in multivalent systems, we consider a case where a bivalent receptor (R – R) binds two monovalent ligands (L). In this case, the rate of formation of the first monovalent species (R – RL) is:

$$\frac{d[R-RL]}{dt} = 2k_{on}[R-R][L] - k_{off}[R-RL]$$

and the subsequent bivalent species (RL – RL) is:

$$\frac{d[RL-RL]}{dt} = k_{on}[R-RL][L] - 2k_{off}[RL-RL]$$

We assume that there is no cooperation between the first and second binding events, so the association and dissociation constants are the same in the two equations above.

In the case where a bivalent ligand ($L-L$) binds to a bivalent receptor ($R-R$) to yield a bivalent receptor-ligand complex ($RL=RL$), we have:

$$\frac{d[RL_L-R]}{dt} = 2k_{on}[R-R][L-L] - k_{off}[RL_L-R]$$

$$\frac{d[RL=RL]}{dt} = k_{on}^{eff}[RL_L-R] - 2k_{off}[RL=RL]$$

where RL_L-R is the partially bound bivalent receptor-ligand complex. We note that the main difference between the two cases is the association phase of the second binding event, where we have $k_{on}^{eff}[RL_L-R]$ instead of $k_{on}[R-RL][L]$. In the bivalent ligand case, the rate of the second reaction is only dependent on the concentration of RL_L-R, and is thus a first-order reaction. (The "2" coefficient in the equations above accounts for both multiple microstates and state symmetry, described later in this section.)

In addition, to fully describe all possible binding events, we consider formation of the doubly monovalent configuration (RL_L-RL_L) in which free ligand competes with the intracomplex ligand in the RL_L-R configuration:

$$\frac{d[RL_L-RL_L]}{dt} = k_{on}[RL_L-R][L-L] - 2k_{off}[RL_L-RL_L]$$

Further, we make a distinction between "inline" and "twisted" bivalent states. With their binding domains numbered "1" and "2" (from N-terminus to C-terminus), bivalent protein receptors and ligands can be written as R_1-R_2 and L_1-L_2, respectively. This yields the inline ($R_1L_1=R_2L_2$) and twisted ($R_1L_2=R_2L_1$) bivalent species (systemization of this multivalent nomenclature is described in *section 3)*. The two bivalent configurations are distinguished by their respective first-order rate constants $k_{on,inline}^{eff}$ and $k_{on,twisted}^{eff}$, resulting in the rate equations:

$$\frac{d[R_1L_1=R_2L_2]}{dt} = k_{on,inline}^{eff}[R_1L_{1\,L_2}-R_2] - 2k_{off}[R_1L_1=R_2L_2]$$

$$\frac{d[R_1L_2=R_2L_1]}{dt} = k_{on,twisted}^{eff}[R_1L_{2\,L_1}-R_2] - 2k_{off}[R_1L_2=R_2L_1]$$

$k_{on,inline}^{eff}$ and $k_{on,twisted}^{eff}$ are equal when both configurations have equal steric

permissibility (e.g., an infinitely long, flexible linker spanning the two binding domains). Our structure-based derivation of these first-order association rate constants is described in *sections 4 and 5*. Application of these treatments to receptor-ligand interactions with valencies higher than two is described in *section 6*.

Finally, to reduce the size of the system, topologically equivalent or 'symmetric' configurations were grouped together. Here, configurations are regarded as symmetric when all of their immediate connections to other microstates (via single binding or dissociation events) are identical. For example, the states $(R_1 L_{1\,L_2} - R_2)$ and $(R_1 - R_2 L_{2\,L_1})$ are symmetric because they derive from the same configuration (e.g., unbound receptor, $R_1 - R_2$) and can yield the same configuration (e.g., $R_1 L_1 = R_2 L_2$) via a synonymous association event (types of associations are described in *section 4*). Absent symmetry considerations, a bivalent ligand and a bivalent receptor would have 4 possible ways of initially associating; with symmetry, there are only 2 topologically distinct ways to associate, leading to the "2" coefficient in the equations above.

3. Modeling combinatorial complexity

3.1 Enumerating all possible binding configurations

In contrast to other models of multivalent interaction, we viewed multivalency as a combinatorial set of binding states interconnected through a network of configurational transitions. Here, explicitly enumerating every unique, non-symmetric binding state was essential to fully capture the complexity of multivalent interactions.

In the case of a bivalent receptor – bivalent ligand interaction, we can readily list all possible configurations: $R - R$, $RL_L - R$, $RL_L - RL_L$ and $RL = RL$. However, in cases of higher valency, the number of configurations increases factorially. If only one ligand binds to one receptor, then the number of states is given by $\sum_{i=1}^{k} P(v_{max}, v)$, where P is the permutation of v_{max}, which is *max(receptor valency, ligand valency)*, and v, which is *min(receptor valency, ligand valency)*.

In addition, further combinatorial binding configurations arise when multiple ligands bind

to a single, multivalent receptor (e.g., $RL_L - RL_L$). Such binding configurations become prominent when, for example, the valency of the receptor is greater than that of the ligand and/or the ligand concentration is high relative to the intracomplex effective concentration. These high-stoichiometry configurations play an additional role as transiently-populated high-energy states that then drive the relaxation to more favorable configurations as the system approaches equilibrium. Consideration of high-stoichiometry states was thus necessary for a complete description of multivalent binding. Without considering the binding order on the receptor, the number of possible states is given by $\sum_{n=1}^{v_{max}} p_v(n)$ where $p_v(n)$ is the partial integer partition of the number of occupied receptor units, *n*, in which the largest part does not exceed *v* (2). Taking receptor binding order into consideration, the number of states can be written as $P(mv,n)\binom{v_{max}}{v_{max}-n}$, where *m* is the number of ligand molecules binding to a receptor and *n* is the number of occupied receptor units. If we do not differentiate between ligand molecules with the same valency, then the previous form is further modified to $\frac{P(mv,n)\binom{v_{max}}{v_{max}-n}}{\prod_{i=1}^{v} num\_val(i)!}$, where *num_val(i)* is a function that counts the number of ligands bound to the receptor with *i* valency.

To handle this combinatorial complexity computationally, we created a system in which the states are treated as an N_{receptor}-long vector: empty receptor units have values of 0 (e.g., empty bivalent receptors are 00 in our computational notation or --/-- in our standardized nomenclature) and bound receptors have values equal to the value of the ordered ligand unit. For example, a fully bound bivalent receptor-bivalent ligand complex is designated $\mathrm{A}_2^1/\mathrm{A}_2^2$ (“inline”) or $\mathrm{A}_2^2/\mathrm{A}_2^1$ (“twisted”): the slash partitions the two receptor binding sites; the ligand identities are given by a capital letter (here only one ligand, so A); the valency of the ligand is given by the immediate subscript (here bivalent); and the bound ligand domain (numbered starting from the N-terminus) is the superscript. Thus, the high-stoichiometry states involving a second bivalent ligand “B” would take the forms $\mathrm{A}_2^1/\mathrm{B}_2^1$, $\mathrm{A}_2^1/\mathrm{B}_2^2$, $\mathrm{A}_2^2/\mathrm{B}_2^1$, and $\mathrm{A}_2^2/\mathrm{B}_2^2$.

3.2 Configurational network of states

We create the configurational network of the multivalent system by considering each

possible configurational transition (i.e., binding or dissociation event) to draw all possible connections among the states.

All the configurational transitions can be collected into an adjacency matrix ***A***, where row position *i* represents the pre-transition states, column position *j* represents the post-transition states, and $\boldsymbol{A_{ij}}$ is the rate constant for transition between two states (an intermolecular or intramolecular association or dissociation rate constant, depending on the nature of the transition). Using matrix ***A***, we can formulate a series of ordinary differential equations that describe the kinetics of binding and dissociation at each node in the network, and macroscopic binding kinetics can be readily determined by summing up the time-dependent concentrations of all the states.

4. Biophysical quantities required to model dynamics in a bivalent configurational network

Modeling bivalent modes of interaction requires an accurate calculation of the first-order, intracomplex rate constant of association, k_{on}^{eff}, that determines the rate of formation of the bivalent species $R_1L_1 = R_2L_2$, and $R_1L_2 = R_2L_1$ (described in *section 3.1*). Here, we define k_{on}^{eff} as the product of the monovalent k_{on} and a measure of the relative intramolecular position of the receptor and ligand binding domains, which takes the form of an effective ligand concentration, $[L_{eff}]$. Calculation of $[L_{eff}]$ requires consideration of the geometric parameters that can both favor and impede intracomplex association: the diameters of the bivalent receptor/ligand binding domains and the structure of the linker between them. Moreover, because the bivalent configurations $R_1L_1 = R_2L_2$, and $R_1L_2 = R_2L_1$ represent two structurally distinct conformations (inline and twisted, respectively), it is essential to independently calculate a $[L_{eff}]$ for each unique intracomplex association event. In contrast to other treatments of effective concentrations that calculate $[L_{eff}]$ assuming a uniform ligand distribution in a confined volume, we created a structure-based determination of $[L_{eff}]$ around a worm-like chain model (3, 4) to yield a set of conditional probability density functions (PDFs) that describe the permissibility, and thus first-order association rates, for all combinations of inline, twisted, proximal, and distal intracomplex binding. For example, Fig. S6 depicts the 12 intracomplex associations that can occur

between a trivalent receptor and ligand.

5. Intramolecular effective concentration model

Calculating the set of $[L_{eff}]$ for all unique, intracomplex association reactions is necessary to determine the k_{on}^{eff} values and to solve the system of differential equations. These values can be approximated by determining what concentration of freely diffusible ligand domains would be equivalent to the concentration of free ligand domains constrained by the binding of its neighboring ligand domains in the same molecule. A common approach to calculate the effective concentration is to assume homogenous distribution of the ligand and receptor free ends in a volume determined by the maximal linker length (r_{max}) to get the concentration of 1 molecule in $\frac{4}{3}\pi r_{max}^3$ volume. Our model refines this approach by taking into account the uneven linker end-to-end distribution.

To calculate the effective concentration in our model, we evaluate the probability of binding between the ligand and receptor and define the condition that they are bound as their coincidence in space. We can formalize this treatment using convolution in which we have two variables L and R, with distributions f_L and f_R, and determine the probability of their being equal. To do this, we introduce the variable Z, where $Z = L - R$ and the desired probability is $p(Z = 0)$. Z is the sum of L and $-R$, so the PDF for Z can be given by the convolution of L and $-R$, or $f_Z(Z) = (f_L \cdot f_{-R})(Z) = \int f_L(V) f_R(V - Z) dV$. Since we are interested only in the case where R and L are equal, we set $Z = 0$, in which case $f_Z(Z = 0) = \int f_L(V) f_R(V) dV$. When we have two models – one with a uniform distribution and one with a linker-derived distribution – we can compare the collision odds in the two models with an odds ratio: $\frac{f_{linker-derived}(Z=0)}{f_{uniform}(Z=0)}$. This odds ratio is our effective concentration divided by that in the uniform model, so we can calculate our $[L_{eff}]$ as follows:

$$[L_{eff}] = \frac{\int_{V=1} f_R(V) \cdot f_L(V) dV}{\int_{V=1} f_R(V) \cdot f_L(uni) dV} \cdot \frac{3}{4\pi r_{max}^3}$$

where V is a unit volume, r_{max} is the maximal end-to-end distance, $f_R(V)$ is the three-dimensional receptor PDF, $f_L(V)$ is the three-dimensional ligand PDF, and $f_L(uni)$ is a uniform ligand PDF. The calculations of $f_R(V)$ and $f_L(V)$ are based on the linker end-to-

end PDF described by

$$P_{linker}(r) \propto 4\pi r^2 \frac{1}{(1-\frac{r^2}{{r_{max}}^2})^{\frac{9}{2}}} e^{\left(-\frac{9}{8l_p}\frac{r_{max}}{(1-\frac{r^2}{{r_{max}}^2})}\right)} \qquad (3, 4)$$

where *r* is the end-to-end distance, r_{max} is the maximal length or contour length, and l_p is the persistence length of the linker.

Using the symmetry of the distribution, the end-to-end distribution probabilities in three dimensions can be calculated using

$$P_{linker}(V) \propto \frac{1}{(1-\frac{(x^2+y^2+z^2)}{{r_{max}}^2})^{\frac{9}{2}}} e^{\left(-\frac{9}{8l_p}\frac{r_{max}}{(1-\frac{(x^2+y^2+z^2)}{{r_{max}}^2})}\right)}$$

where x, y, z are the coordinates describing a point V in space. The integral of the probability is normalized to 1 to obtain the linker PDF.

To acquire the odds ratios, a second normalization is needed. Here, the $f_{linker-derived}$ and the $f_{uniform}$ PDFs have different ranges and are normalized by integrating over unit volume while maintaining their $\frac{r_{max}}{l_p}$ ratios.

Moreover, to account for the distance between the receptor/ligand active sites and the binding point of the linker, we created a joint PDF of the linker and a hinged rod representing this distance. Here, rods representing the receptor/ligand binding domains were described with context-dependent persistence lengths (l_p). This added feature allows for more accurate spatial description of $f_R(V)$ and $f_L(V)$ in instances where the receptors and/or ligands undergo conformational changes upon binding that alter their flexibilities.

6. Extension of bivalent framework to higher valency interactions

The framework described above for bivalent receptor-ligand interactions is readily extensible to higher valency interactions. For example, for trivalent interactions, the

enumeration of the binding configurations entails an added increment to the nomenclature described in *section 3.1*. A trivalent receptor (--/--/--) interacting with a trivalent ligand can adopt six fully bound configurations: $A_3^1/A_3^2/A_3^3$, $A_3^1/A_3^3/A_3^2$, $A_3^2/A_3^1/A_3^3$, $A_3^2/A_3^3/A_3^1$, $A_3^3/A_3^1/A_3^2$, $A_3^3/A_3^2/A_3^1$. Of these, $A_3^1/A_3^3/A_3^2 \equiv A_3^2/A_3^1/A_3^3$ and $A_3^2/A_3^3/A_3^1 \equiv A_3^3/A_3^1/A_3^2$. Further, trivalent receptors can form high-stoichiometry configurations with a second or third ligand ("B" or "C") to yield such states as $A_3^1/B_3^2/A_3^3$ and $A_3^3/B_3^2/C_3^1$.

Additionally, each incremental increase in receptor and/or ligand valency is treated as an additional linker and unit segments. The bivalent receptor $f_R(V)$ PDF itself consists of two unit PDFs and one linker PDF

$$f_R(V) = \int_{V_2|V_1,V_0} \left(f_{unit}(V_3|V_2,V_1,V_0) \cdot \int_{V_1|V_0} f_{linker}(V_2|V_1,V_0) \cdot f_{unit}(V_1|V_0) \right)$$

where V_0 and V_3 represent the position of the two binding domains and V_1 and V_2 are the positions of the linker-protein joints. Each increment of valency is thus represented as an added linker and unit term to the chain.

7. Synthesis of the system of differential equations

The calculated $[L_{eff}]$ values yield all of the intramolecular k_{on}^{eff} values, and k_{off}^{eff} is the same as the monovalent k_{off}, so the differential equation system is fully parameterized. The system of differential equations can be generated by the following equation $\frac{\boldsymbol{ds}}{dt} = \boldsymbol{A}^T\boldsymbol{s} - \sum_{i=1}^{n}(\boldsymbol{A_{ij}} \cdot \boldsymbol{s_i})$ where $\boldsymbol{s}$ is the column vector of state concentrations, $\frac{\boldsymbol{ds}}{dt}$ is the derivative of the state concentrations, and $(\sum_{i=1}^{n} \boldsymbol{A_{ij}})$ is the row sum of the adjacency matrix $\boldsymbol{A}$ (Fig. S1D, S2, S3). The time-dependent concentrations for each state can be calculated numerically using conventional ordinary differential equation solvers, and we used the ode15s solver in MATLAB for this purpose (5). The full configurational network was further visualized using Cytoscape (https://cytoscape.org/).

8. Relating theoretical simulations to surface plasmon resonance experiments

In SPR, binding occurs on a chip surface and within a matrix of a defined thickness (1). This results in a thin, three-dimensional interaction volume. For the model to provide an

experimentally grounded description of multivalency, the following treatments were applied. First, the molecular weights of the receptor and ligand and the dimensions of the flow cell were used to convert molar units of concentration into a Biacore-mimicking resonance unit (RU), where 1 RU equals 1 pg of ligand mass bound to a 1 mm^2 area. Second, instances of high receptor concentration within the chip matrix and/or long linkages between binding domains of the ligand, crosslinking can occur in which a single multivalent ligand engages two or more receptors simultaneously. This effect of crosslinking confounds monovalent interactions with apparent multivalent kinetics. To assess and adjust for crosslinking, nearest-neighbor distributions (nnds) of randomly deposited, covalently tethered receptors in a chip matrix were simulated for a given concentration of receptor. The nnds were simulated as a continuum of bivalent receptors with rigid linkages equal to each of the nnd histogram bins (Fig. S12B). Summation of the nnd simulations yields crosslinking-corrected kinetics. Third, simulated mass-transport limitations are applied to account for instances of multivalent ligand re-binding the chip surface following dissociation.

Recombinant receptor-ligand cloning, expression, and purification

The receptor domain used in experimental binding studies was obtained from the C-terminal SH3 domain of the human adaptor protein Gads (6). A peptide sequence (SPAPSIDRSTKPPL) derived from Gads cognate ligand, SLP-76, was used as the ligand binding domain (6). DNA sequences of the receptors and ligands were synthesized as gBlocks (Integrated DNA Technologies) and were introduced into His_6-tag-encoding pET28a (SH3 receptors; Novagen) or MBP tag-encoding pMal-c5x (SLP-76 peptide; New England BioLabs) expression vectors with restriction enzyme DNA cloning methods. Multiple cloning sites in the expression vectors were used to create modular repeats of indicated valencies separated by peptide linkers. The interdomain polypeptide linkers used for both receptors and ligands where designed as either “short” or “long”, “flexible” (i.e., random coil) or “rigid” (i.e., alpha-helical). The following sequences were used: “short flexible”, GSTSGDNSNSGGSGNSGGSGGN; “long flexible”, GSTGGDNSNSGGSGNSG GSGGNSGSTSGDNSNSGGSGNSGGASGN; “short rigid” SPAEAAAKEAAAKEAAAKE

AAAKAPS; and “long rigid”, SPAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEA
AAKAPSGPAPSIDASTAPPLGSPAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAK
EAAAKAPS. For recombinant protein production, *Escherichia coli* BL21 cells were transformed with receptor and ligand expression vectors. Protein expression was induced with 0.5 mM IPTG followed by overnight shaking and incubation at 17°C. Proteins were purified by Talon affinity (His_6-tagged receptors; Clontech) or Amylose resin (MBP-tagged ligands; NEB), followed by size exclusion chromatography (Hiload Superdex S200; GE Life Sciences) and buffer exchange by dialysis or desalting column (ThermoFisher). For use in immobilization for surface plasmon resonance, pET28a::receptor plasmids were additionally modified with insertion of an N-terminal biotinylation tag (Avidity AviTag™). AviTagged receptors were biotinylated by co-transformation of BL21 with GST-BirA. Biotinylated receptors were purified as described above.

Surface plasmon resonance

1. Instrumentation

Surface plasmon resonance experiments were performed at 25˚C on a Biacore S200 (GE Life Sciences). Running buffer HBS-EP+ (10 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.05% Tween-20, pH 7.4) was used for all immobilizations and binding experiments.

2. Immobilization of biotinylated receptors

Biacore CM5 sensor chip flow cells were preconditioned with two 12 s injections of each of 50 mM NaOH, 20 mM HCl, and 0.1% w/v SDS. NeutrAvidin immobilized surfaces were generated on all flow cells. 100 µg/ml NeutrAvidin (ThermoFisher) was prepared in 10 mM sodium acetate, pH 4.5. The amine coupling kit (GE Life Sciences) was used to activate the CM5 chip surfaces with the immobilization software protocol set to immobilize 2500 RU NeutrAvidin on each of the four flow cells. Following NeutrAvidin immobilization, flow cells were quenched with 100 mM ethanolamine in 150 mM borate, pH 8.5.

3. Kinetic assays

Kinetic assays were performed in HBS-EP+ running buffer. Biotinylated receptors were

prepared at concentrations ranging from 50-500 ng/mL and immobilized at RU values indicated for each figure (5-75 RU). To reduce experimental variation, monovalent, bivalent, and trivalent receptors were generally immobilized in series to adjacent flow cells (FC), FC2, FC3, and FC4, respectively. NeutrAvidin-coated FC1 was used as the reference flow cell and experiments were conducted with buffer blanks run at least in triplicate for double referencing. Ligands were prepared in 10 mM stock solutions and serially diluted across desired concentration ranges in HBS-EP+. Concentration series duplicates were performed to assess cycle-to-cycle variation. To minimize the occurrence of injection spikes, experiments were performed at a faster flow rate, 75 µL/min.

4. SPR data analysis

Experimental data for the monovalent receptor-ligand interaction were fit with a 1:1 Langmuir model for kinetic fits, and the equilibrium analysis model for RU_{max} and equilibrium dissociation constant fits.

Supplementary References

[1] N.J. deMol, M.J.E. Fischer, “Surface plasmon resonance: a general introduction” in *Surface Plasmon Resonance: Methods and Protocols*, N.J. deMol, M.J.E. Fischer, Eds. (Humana Press, 2010), pp. 1-14.

[2] R.P. Stanley, *Enumerative Combinatorics* (Cambridge University Press, 1999).

[3] D. Thirumalai, B.Y. Ha, “Statistical mechanics of semiflexible chains: A meanfield variational approach” in *Theoretical and Mathematical Models in Polymer Research*, A. Grosberg, Ed. (Academic Press, 1998), pp. 1–35.

[4] U.G. Zinth. “End-to-End Distance Distribution and Intra-Chain Diffusion in Unfolded Polypeptide Chains.” Technische Universität München, Germany (2014).

[5] L.F. Shampine, M.W. Reichelt, The MATLAB ODE Suite. *SIAM Journal on Scientific Computing* Vol. 18, pp. 1–22 (1997).

[6] Q. Liu et al., Structural basis for specific binding of the Gads SH3 domain to an RxxK motif-containing SLP-76 peptide: a novel mode of peptide recognition. *Mol. Cell* **11**:471–81 (2003).

[7] H.M. Kim et al., Crystal structure of the BAFF-BAFF-R complex and its implications for receptor activation. *Nat. Struct. Mol. Biol.* **10**, 342–348 (2003).

[8] S. Radaev, S. Motyka, W.-H. Fridman, C. Sautes-Fridman, P.D. Sun, The structure of a human type III Fcγ receptor in complex with Fc. *J. Biol. Chem.* **276**, 16469–16477 (2001).

[9] E.S. Day et al., Selectivity of BAFF/BLyS and APRIL for Binding to the TNF Family Receptors. *Biochemistry* **44**, 1919–1931 (2005).

[10] S.Y. Breusegem, R.M. Clegg, F.G. Loontiens, Base-sequence specificity of Hoechst 33258 and DAPI binding to five $(A/T)_4$ DNA sites with kinetic evidence for more than one high-affinity Hoechst 33258-AATT complex. *J. Mol. Biol.* **315**, 1049–1061 (2002).

[11] C.G. Baumann, S.B. Smith, V.A. Bloomfield, C. Bustamante, Ionic effects on the elasticity of single DNA molecules. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 6185–6190 (1997).

[12] H. Lee, R.M. Venable, A.D. Jr. Mackerell, R.W. Pastor, Molecular dynamics studies of polyethylene oxide and polyethylene glycol: hydrodynamic radius and shape anisotropy. *Biophys. J.* **95**, 1590–1599 (2008).

[13] Y. Liu et al., Designed Compounds for Recognition of 10 Base Pairs of DNA with Two AT Binding Sites. *J. Am. Chem. Soc.* **134**, 5290–5299 (2012).

A. Enumerate all unique binding configurations

I. All possible configurations:

i. Empty space = *receptor_valency* = 2 =␣ ␣

ii. All possible ligand units

1. A maximum of 2 (*receptor_valency*) different ligands can bind to a receptor
2. Each receptor has 2 (*ligand_valency*) different units
3. Therefore, we have 2 (*receptor_valency*) * 2 (*ligand_valency*) different ligand units

iii. The possible values on a receptor unit (␣) are the list of:
[*receptor_valency* * *empty, receptor_valency* * *ligand_valency*]

iv. For the combinations we need to choose *receptor_valency* (␣␣) values from the list; e.g., bivalent-bivalent: choose two from [empty, empty, ligand1_unit1, ligand1_unit2, ligand2_unit1, ligand2_unit2]. Thus, [ligand1_unit1, ligand1_unit2] = A_2^1/A_2^2

II. All actual states:

i. Identify redundant state names $A_2^1/A_2^2 = B_2^1/B_2^2$

ii. Identify symmetric configurations

iii. Use unique states and compensate for the multiplicity by a multiplier

B. Identify all possible association/dissociation transitions between configurations

I. Compare each state with each other (imagine an adjacency matrix)

i. If the difference is one unit, then they are connected

ii. If a state has 1 more empty unit than the current state, then it can transform to it by association

iii. If a state has 1 fewer empty unit than the current state, then it can transform to it by dissociation

iv. For mass balance, the multiplicity of the states are applied to the dissociation and the association

II. Identifying the type of the association

i. If a new ligand binds to a receptor (not previously bound ligand with new unit), then it is assigned monovalent association kinetics

ii. If a new unit of a previously bound ligand binds to the receptor, then it has multivalent kinetics described by multivalent binding kinetics (combination of effective concentrations and monovalent kinetics)

III. Identifying the dissociation values

i. All the dissociation events have the same monovalent value (k_{off}) over all the states. This assumes multivalent binding relies on steric accessiblility and does not create excessive strain on the configuration or alter the binding interface.

C. Calculate the effective concentrations for each unique "intra-complex" association

I. Multivalent binding is equivalent to monovalent binding with effective concentration instead of bulk ligand concentration.

II. Effective concentration: the concentration of uniformly distributed, soluble ligand that binds with an equal probability to the empty unit of the multivalent receptor as the other unit of the already bound ligand, given by the odds ratio:

$$[L_{eff}] = \frac{\int_{V=1} f_R(V) \cdot f_L(V) dV}{\int_{V=1} f_R(V) \cdot f_L(uni) dV} \cdot \frac{3}{4\pi r_{max}^3}$$

where V is a unit volume, r_{max} is the maximal end-to-end distance, $f_R(V)$ is the three dimensional receptor PDF, $f_L(V)$ is the ligand three dimensional PDF and $f_L(uni)$ is the uniform ligand PDF.

III. Generating PDF for the linker and for the receptor

i. The linker PDF can be approximated by the end-to-end distance PDF characteristic of the linker (like peptide chain end-to-end distance PDF)

ii. Creating the linked PDF of the linker and the receptor or ligand unit (in case of trivalent: unit-linker-unit-linker-unit linked PDF)

D. Solve the system of differential equations

I. Creating the differential equation system:

i. The association and dissociation are summed up into an adjacency matrix where association has positive signs and dissociation has negative signs

ii. The differential equation system (E) is the product of the concentration vector of the states (C) and the transpose of the adjacency matrix (A^T) minus the product of the concentration vector of the states (C) and the the adjacency matrix (A)

$$A = \begin{bmatrix} \square & \cdots & kon \\ \vdots & \ddots & \vdots \\ koff & \cdots & \square \end{bmatrix}, C = \begin{pmatrix} C_{state1} \\ C_{state2} \\ \cdots \\ C_{staten} \end{pmatrix}, E = \begin{pmatrix} \frac{dCstate1}{dt} \\ \frac{dCstate2}{dt} \\ \cdots \\ \frac{dCstaten}{dt} \end{pmatrix} \quad E = CA^T - CA$$

II. The differential equations can be calculated by a differential equation solver to yield the results

III. Obtain values for each concentration

Fig. S1. The zero-fit model of multivalency is based on a four-part design scheme. Here, as an example, the computational process for a bivalent receptor-bivalent ligand system is described. **(A)** For a set of user-specified receptor-ligand valencies, all possible and unique multivalent binding configurations are enumerated regardless of their steric permissibility. (*I*) The number of ways in which a receptor may be engaged by ligand is described in *section 3.1*. (*II*) To simplify the enumeration, redundant state names are filtered out (*i*), and symmetric configurations are identified (*ii*) and combined with a multiplier (*iii*) to ensure mass balance of the ensemble. Each unique configuration is represented as a network node. **(B)** All possible association and dissociation transitions between the configurations are identified and represented as connections (edges) between pairs of network nodes. (*I*) Adjacency matrices are created in the form of k_{on} and k_{off} tables. (*II*) Values in the k_{on} table represent either monovalent or multivalent association kinetics. (*III*) All dissociation transitions are assigned monovalent kinetics in the k_{off} table. **(C)** Effective ligand concentrations are calculated for each intracomplex association. (*I*) A multivalent association constant is equal to the monovalent constant (k_{on}) multiplied by the effective concentration of the intracomplex free ligand $k_{on}^{eff}=k_{on}\cdot[L_{eff}]$. (*II*) Here, we define $[L_{eff}]$ as the concentration of soluble, monovalent ligand required to bind receptor with a probability equal to the intracomplex association. The probability of an intracomplex association is the odds ratio between an intracomplex receptor and ligand binding event and one involving a uniform distribution of monovalent ligand. **(D)** Combined $[L_{eff}]$ PDFs and a system of differential rate equations yield time-resolved association and dissociation of a multivalent configurational ensemble.

	x0x0	a1x0	a2x0	a1a2	a1b1	a1b2	a2a1	a2b1
x0x0		2*kon	2*kon					
a1x0	koff			kons11	kon	kon		
a2x0	koff				kon		konr11	kon
a1a2		2*koff						
a1b1		koff	koff					
a1b2		2*koff						
a2a1			2*koff					
a2b1			2*koff					

Fig. S2. Adjacency matrix for a bivalent receptor-bivalent ligand network, specifying possible configurational transitions with assignment of monovalent or multivalent kinetic rate constants. k_{on} and k_{off} are monovalent kinetic rates constants. k_{ons11} and k_{onr11} are the effective rate constants for the inline and twisted bivalent configurations, respectively.

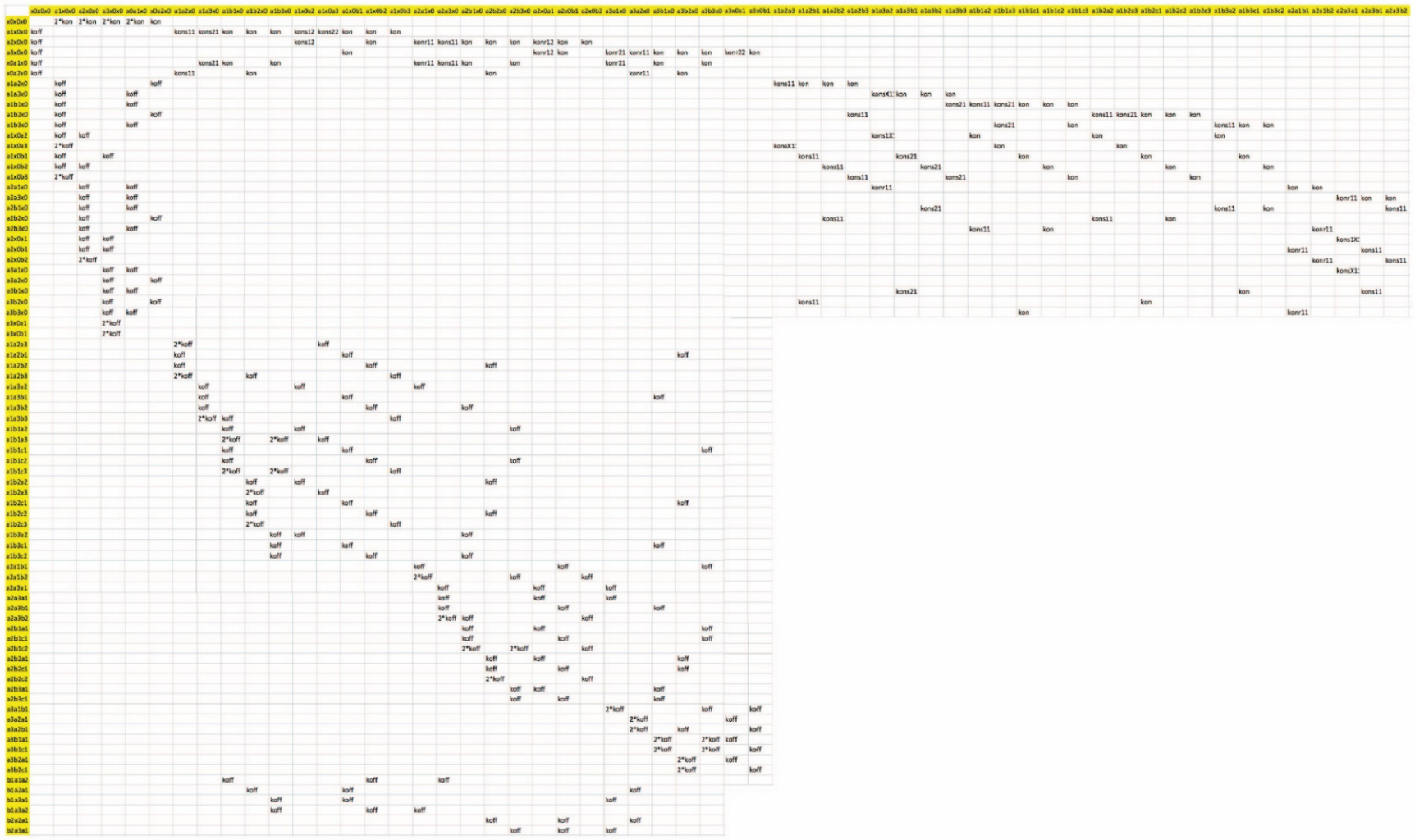
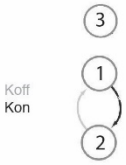


Fig. S3. Adjacency matrix for a trivalent receptor-trivalent ligand network, specifying possible configurational transitions with assignment of monovalent or multivalent kinetic rate constants. k_{on} and k_{off} are monovalent kinetic rates constants. k_{ons}** and k_{onr}** are the effective rate constants for the inline and twisted bivalent configurations, respectively.

A Monovalent - Monovalent

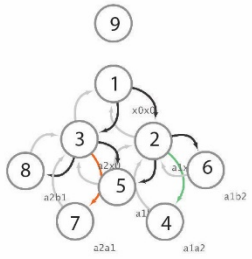
```
dydt(1)=+1*Koff*y(2)-1*Kons00*y(1)*y(3);
dydt(3)=km*L-km*y(3)+1*Koff*y(2)-1*Kons00*y(1)*y(3);

dydt(2)=+1*Kons00*y(1)*y(3)-1*Koff*y(2);
```



B Bivalent - Bivalent

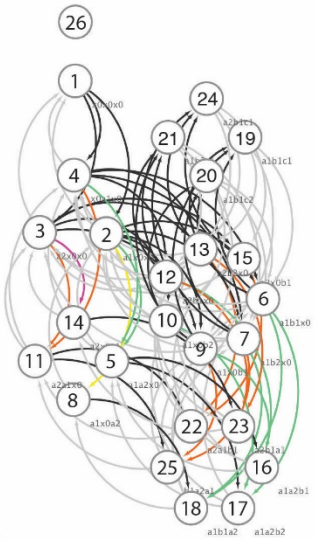
```
dydt(1)=+1*Koff*y(2)+1*Koff*y(3)-2*Kons00*y(1)*y(9)-2*Kons00*y(1)*y(9);
dydt(9)=km*L-km*y(9)+1*Koff*y(2)+1*Koff*y(3)-2*Kons00*y(1)*y(9)-2*Kons00*y(1)*y(9);
dydt(2)=+2*Kons00*y(1)*y(9)+2*Koff*y(4)+1*Koff*y(5)+2*Koff*y(6)-1*Koff*y(2)-1*Kons11*y(2)-
        1*Kons00*y(2)*y(9)-1*Kons00*y(2)*y(9);
dydt(3)=+2*Kons00*y(1)*y(9)+1*Koff*y(5)+2*Koff*y(7)+2*Koff*y(8)-1*Koff*y(3)-1*Kons00*y(3)*y(9)-
        1*Konr11*y(3)-1*Kons00*y(3)*y(9);
dydt(4)=+1*Kons11*y(2)-2*Koff*y(4);
dydt(5)=+1*Kons00*y(2)*y(9)+1*Kons00*y(3)*y(9)-1*Koff*y(5)-1*Koff*y(5);
dydt(6)=+1*Kons00*y(2)*y(9)-2*Koff*y(6);
dydt(7)=+1*Konr11*y(3)-2*Koff*y(7);
dydt(8)=+1*Kons00*y(3)*y(9)-2*Koff*y(8);
```



C Trivalent - Bivalent

```
dydt(1)=+1*Koff*y(2)+1*Koff*y(3)+1*Koff*y(4)-2*Kons00*y(1)*y(26)-2*Kons00*y(1)*y(26)-2*Kons00*y(1)*y(26);
dydt(26)=km*L-km*y(26)+1*Koff*y(2)+1*Koff*y(3)+1*Koff*y(4)-2*Kons00*y(1)*y(26)-2*Kons00*y(1)*y(26)-
        2*Kons00*y(1)*y(26);

dydt(2)=+2*Kons00*y(1)*y(26)+1*Koff*y(5)+1*Koff*y(6)+1*Koff*y(7)+2*Koff*y(8)+1*Koff*y(9)+1*Koff*y(10)-
        1*Koff*y(2)-1*Kons11*y(2)-1*Kons00*y(2)*y(26)-1*Kons00*y(2)*y(26)-1*Kons12*y(2)-1*Kons00*
        y(2)*y(26)-1*Kons00*y(2)*y(26);
dydt(3)=+2*Kons00*y(1)*y(26)+1*Koff*y(9)+1*Koff*y(11)+1*Koff*y(12)+1*Koff*y(13)+2*Koff*y(14)+1*Koff*y(15)-
        1*Koff*y(3)-1*Kons00*y(3)*y(26)-1*Konr11*y(3)-1*Kons00*y(3)*y(26)-1*Kons00*y(3)*y(26)-1*Konr12*
        y(3)-1*Kons00*y(3)*y(26);
dydt(4)=+2*Kons00*y(1)*y(26)+1*Koff*y(5)+1*Koff*y(6)+1*Koff*y(7)+1*Koff*y(11)+1*Koff*y(12)+1*Koff*y(13)-
        1*Koff*y(4)-1*Kons11*y(4)-1*Kons00*y(4)*y(26)-1*Kons00*y(4)*y(26)-1*Konr11*y(4)-1*Kons00*y(4)*
        y(26)-1*Kons00*y(4)*y(26);
dydt(5)=+1*Kons11*y(2)+1*Kons11*y(4)+1*Koff*y(16)+1*Koff*y(17)-1*Koff*y(5)-1*Koff*y(5)-1*Kons00*y(5)*
        y(26)-1*Kons00*y(5)*y(26);
dydt(6)=+1*Kons00*y(2)*y(26)+1*Kons00*y(4)*y(26)+1*Koff*y(17)+1*Koff*y(18)+1*Koff*y(19)+1*Koff*y(20)-
        1*Koff*y(6)-1*Koff*y(6)-1*Kons11*y(6)-1*Kons11*y(6)-1*Kons00*y(6)*y(26)-1*Kons00*y(6)*y(26);
dydt(7)=+1*Kons00*y(2)*y(26)+1*Kons00*y(4)*y(26)+1*Koff*y(18)+1*Koff*y(20)+1*Koff*y(21)+1*Koff*y(25)-
        1*Koff*y(7)-1*Koff*y(7)-1*Kons11*y(7)-1*Kons00*y(7)*y(26)-1*Kons00*y(7)*y(26)-1*Konr11*y(7);
dydt(8)=+1*Kons12*y(2)+1*Koff*y(18)-2*Koff*y(8)-1*Kons00*y(8)*y(26);
dydt(9)=+1*Kons00*y(2)*y(26)+1*Kons00*y(3)*y(26)+1*Koff*y(16)+1*Koff*y(19)+1*Koff*y(21)+1*Koff*y(25)-
        1*Koff*y(9)-1*Koff*y(9)-1*Kons11*y(9)-1*Kons00*y(9)*y(26)-1*Kons00*y(9)*y(26)-1*Konr11*y(9);
dydt(10)=+1*Kons00*y(2)*y(26)+1*Koff*y(17)+1*Koff*y(20)-1*Koff*y(10)-1*Kons11*y(10)-1*Kons00*y(10)*y(26);
dydt(11)=+1*Konr11*y(3)+1*Konr11*y(4)+1*Koff*y(22)+1*Koff*y(25)-1*Koff*y(11)-1*Koff*y(11)-1*Kons00*y(11)*
        y(26)-1*Kons00*y(11)*y(26);
dydt(12)=+1*Kons00*y(3)*y(26)+1*Kons00*y(4)*y(26)+1*Koff*y(16)+1*Koff*y(21)+1*Koff*y(23)+1*Koff*y(24)-
        1*Koff*y(12)-1*Koff*y(12)-1*Kons11*y(12)-1*Kons00*y(12)*y(26)-1*Konr11*y(12)-1*Kons00*y(12)*y(26),
dydt(13)=+1*Kons00*y(3)*y(26)+1*Kons00*y(4)*y(26)+1*Koff*y(19)+1*Koff*y(22)+1*Koff*y(23)+1*Koff*y(24)-
        1*Koff*y(13)-1*Koff*y(13)-1*Kons00*y(13)*y(26)-1*Konr11*y(13)-1*Konr11*y(13)-1*Kons00*y(13)*y(26);
dydt(14)=+1*Konr12*y(3)+1*Koff*y(23)-2*Koff*y(14)-1*Kons00*y(14)*y(26);
dydt(15)=+1*Kons00*y(3)*y(26)+1*Koff*y(22)+1*Koff*y(24)-1*Koff*y(15)-1*Konr11*y(15)-1*Kons00*y(15)*y(26);
dydt(16)=+1*Kons00*y(5)*y(26)+1*Kons11*y(9)+1*Kons11*y(12)-1*Koff*y(16)-1*Koff*y(16)-1*Koff*y(16);
dydt(17)=+1*Kons00*y(5)*y(26)+1*Kons11*y(6)+1*Kons11*y(10)-1*Koff*y(17)-1*Koff*y(17)-1*Koff*y(17);
dydt(18)=+1*Kons11*y(6)+1*Kons11*y(7)+1*Kons00*y(8)*y(26)-1*Koff*y(18)-1*Koff*y(18)-1*Koff*y(18);
dydt(19)=+1*Kons00*y(6)*y(26)+1*Kons00*y(9)*y(26)+1*Kons00*y(13)*y(26)-1*Koff*y(19)-1*Koff*y(19)-
        1*Koff*y(19);
dydt(20)=+1*Kons00*y(6)*y(26)+1*Kons00*y(7)*y(26)+1*Kons00*y(10)*y(26)-1*Koff*y(20)-1*Koff*y(20)-
        1*Koff*y(20);
dydt(21)=+1*Kons00*y(7)*y(26)+1*Kons00*y(9)*y(26)+1*Kons00*y(12)*y(26)-1*Koff*y(21)-1*Koff*y(21)-
        1*Koff*y(21);
dydt(22)=+1*Kons00*y(11)*y(26)+1*Konr11*y(13)+1*Konr11*y(15)-1*Koff*y(22)-1*Koff*y(22)-1*Koff*y(22);
dydt(23)=+1*Konr11*y(12)+1*Konr11*y(13)+1*Kons00*y(14)*y(26)-1*Koff*y(23)-1*Koff*y(23)-1*Koff*y(23);
dydt(24)=+1*Kons00*y(12)*y(26)+1*Kons00*y(13)*y(26)+1*Kons00*y(15)*y(26)-1*Koff*y(24)-1*Koff*y(24)-
        1*Koff*y(24);
dydt(25)=+1*Konr11*y(7)+1*Konr11*y(9)+1*Kons00*y(11)*y(26)-1*Koff*y(25)-1*Koff*y(25)-1*Koff*y(25);
```



Koff
Kon
Konr11=Kon*(N*R0*(M-1)+Ceffr11);
Kons11=Kon*(N*R0*(M-1)+Ceffs11);

Koff
Kon
Konr11=Kon*(N*R0*(M-1)+Ceffr11);
Konr12=Kon*(N*R0*(M-1)+Ceffr12);
Kons11=Kon*(N*R0*(M-1)+Ceffs11);
Kons12=Kon*(N*R0*(M-1)+Ceffs12);

Fig. S4. Systems of differential rate equations dictate the kinetics and equilibria of multivalent receptor-ligand ensembles. Differential equations, rate constants, and

network ensemble are shown for **(A)** monovalent-monovalent, **(B)** bivalent-bivalent, and **(C)** trivalent-bivalent interactions. Network maps show each unique and non-symmetric configuration as nodes (numbered according to the set of differential equations). Configurational transitions are indicated by edges between nodes. k_{on} (colored black), k_{off} (colored gray), and effective k_{on} values (multi-colored) are shown.

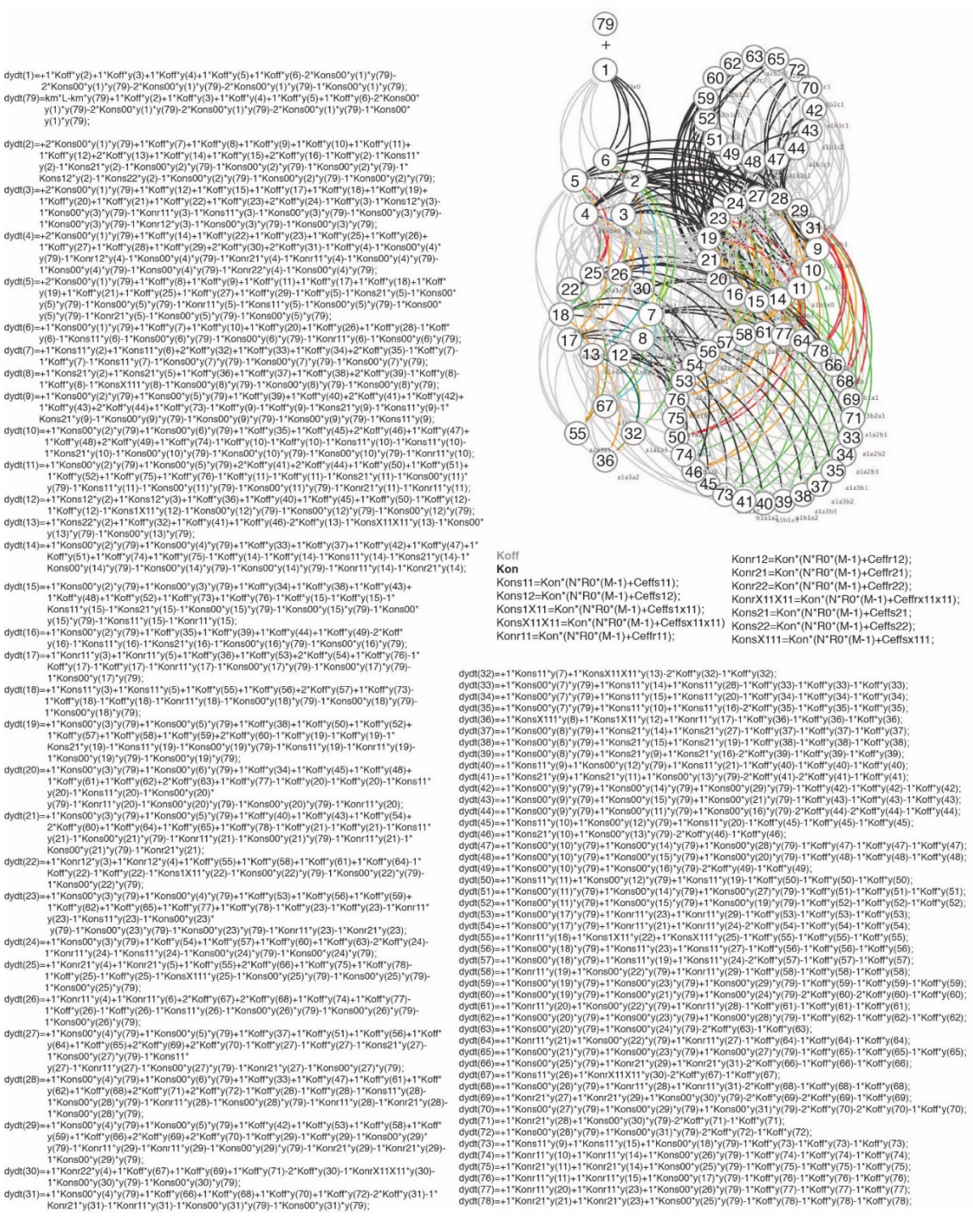


Fig. S5. Differential equations, rate constants, and network ensemble are shown for a trivalent receptor-trivalent ligand (see also Fig. S4). The network map shows each unique and non-symmetric configuration as nodes (numbered according to the set of differential equations). Configurational transitions are indicated by edges between nodes. k_{on} (colored black), k_{off} (colored gray), and effective k_{on} values (multi-colored) are shown.

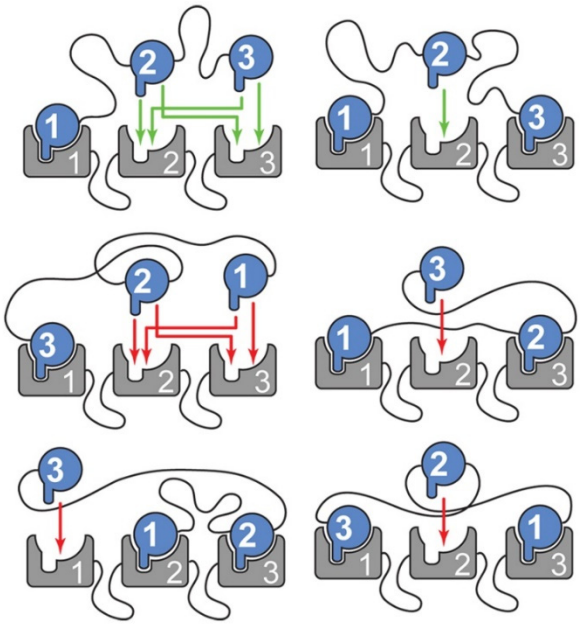


Fig. S6. Trivalent ligand-receptor system involves 12 intramolecular rate constants of association that specify the formation of five inline (green) and seven twisted (red) types of binding events.

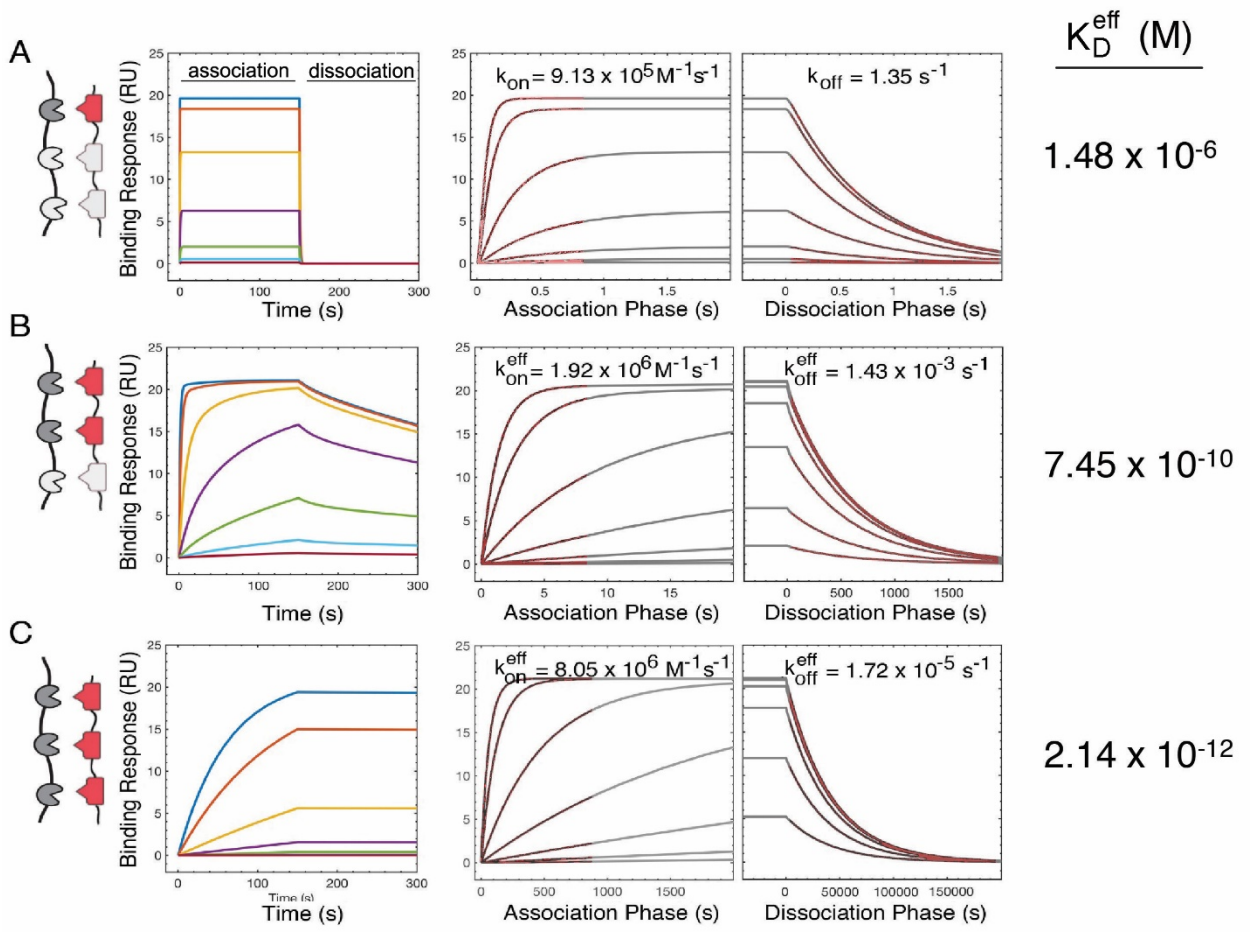


Fig. S7. Simulations of multivalent interactions allow sampling of molecular parameters and timescales that can be difficult or impractical to perform experimentally. **(A)** Monovalent-monovalent, **(B)** bivalent-bivalent, and **(C)** trivalent-trivalent receptor-ligand simulations are shown. Simulations are performed with the same monovalent kinetic input parameters (k_{on} = 913,000 $M^{-1}s^{-1}$ and k_{off} = 1.35 s^{-1}). Association and dissociation phases are varied by orders of magnitude for the monovalent (on = 2 sec, off = 2 sec), bivalent (on = 20 sec, off = 2000 sec), and trivalent (on = 2000 sec, off = 200,000 sec). Effective rate constants of association (k_{on}^{eff}) and dissociation (k_{off}^{eff}) were determined by fitting each phase to mono-exponential, 1:1 Langmuir kinetics.

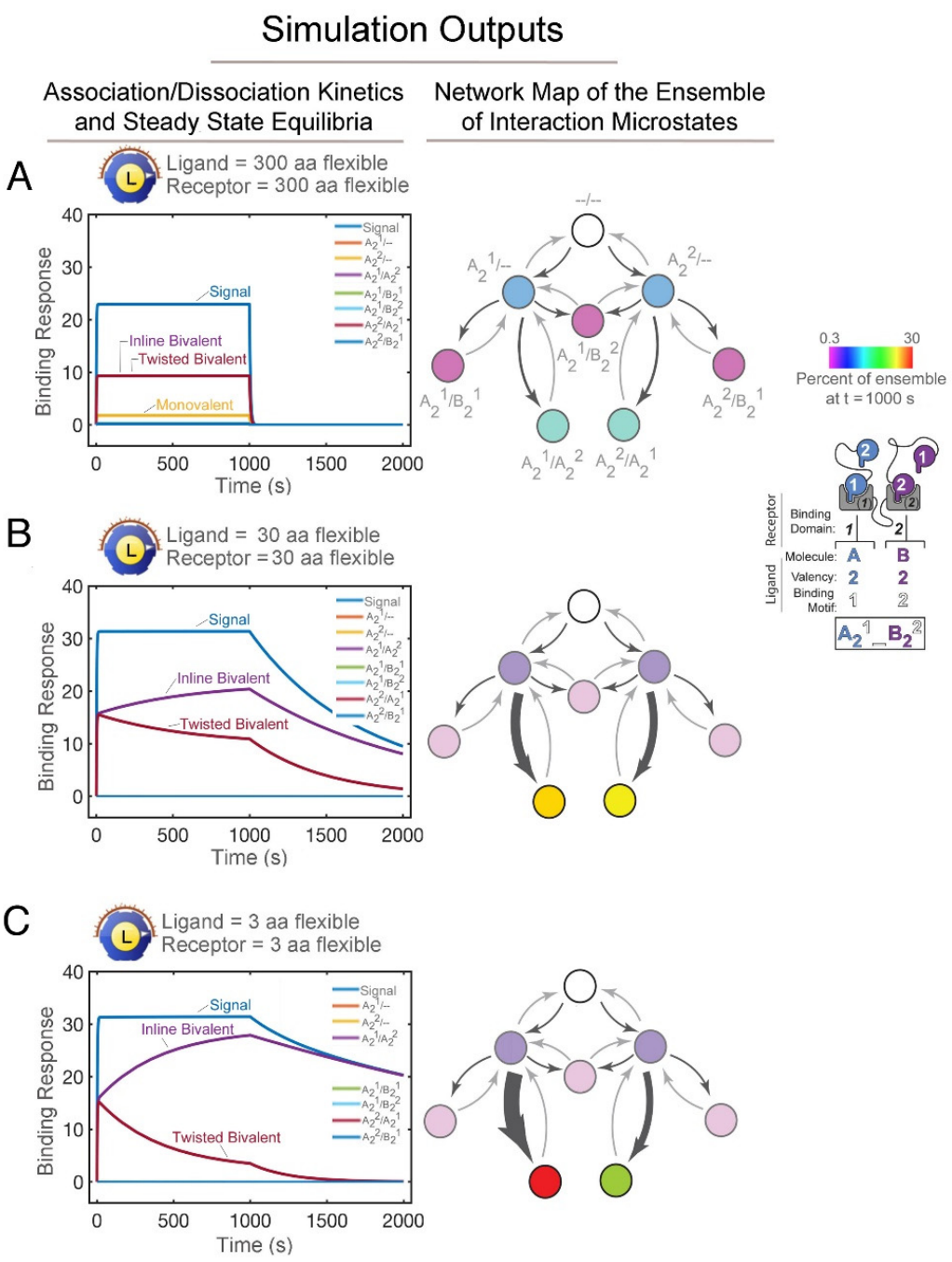


Fig. S8. Simulations performed on a bivalent ligand-receptor interaction with k_{on} = 10^6 M^{-1} s^{-1}, k_{off} = 1 s^{-1}, and varying linkages: (A) 300, (B) 30, and (C) 3 amino acid flexible linkers. Simulated SPR sensorgrams (left) with 1000 sec association and dissociation phases, and network maps of the microstate ensembles (right).

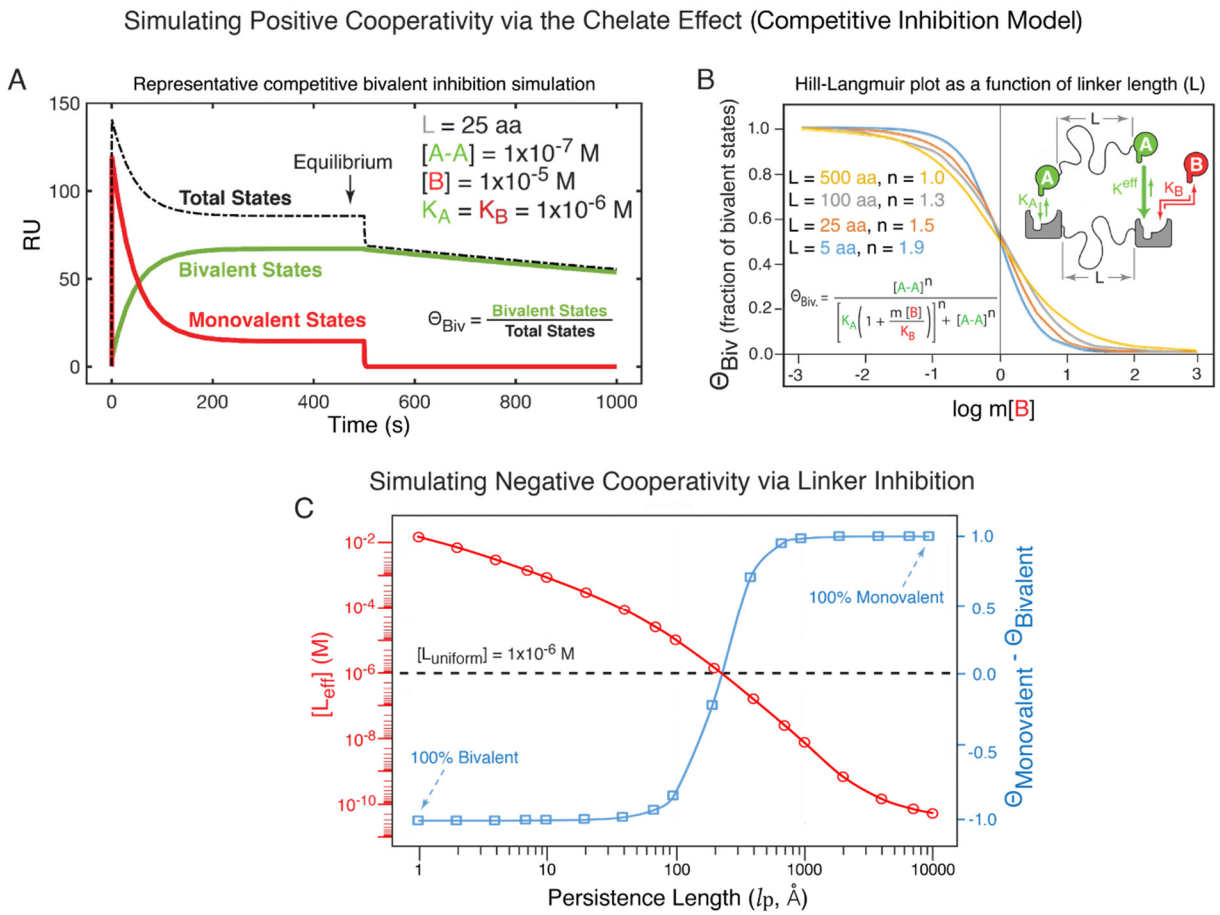


Fig. S9. Manifestation of positive and negative cooperativity in simulated multivalent interactions. (A) Elaboration of the multivalency model enables competitive dissociation of a bivalent receptor-bivalent ligand ("A-A") interaction with a monovalent ligand ("B"). The fraction of bivalent states among the total states is measured as a function of ligand "B" concentration. K_A and K_B refer to the monovalent equilibrium dissociation constants for A and B, respectively. (B) Competitive dissociation curves are simulated for a range of bivalent receptor-bivalent ligand flexible linker lengths (500, 100, 25, and 5 amino acids). Curves are fit with a cooperative inhibition equation to obtain the parameters "m" (a concentration normalization coefficient to align the midpoints of the curves) and "n" (a Hill-like coefficient). Here, with shortening of the linkers (i.e., increasing $[L_{eff}]$), "n" approaches 2, the maximum valency, as the system displays stronger cooperativity between the two ligand "A" binding domains. (C) Negative cooperativity in multivalent interactions arises through linker inhibition. Here, bivalent interactions are simulated for receptor-ligand pairs with linkers of 10 amino acids (the ligand) and 35 amino acids (the receptor). The persistence length (l_p) of both linkers is progressively increased from 1 Å (highly flexible) to 10,000 Å (highly rigid). $[L_{eff}]$ and the fraction of bivalent species are plotted at each l_p sampled. The uniform ligand concentration ($1x10^{-6}$ M) is shown for comparison. Negative cooperativity arising from linker-driven binding events occurs when $[L_{eff}]$ is lower than $[L_{uniform}]$. This is commonly represented as a "cooperativity factor" (CF) < 1, where CF = $[L_{eff}]/[L_{uniform}]$.

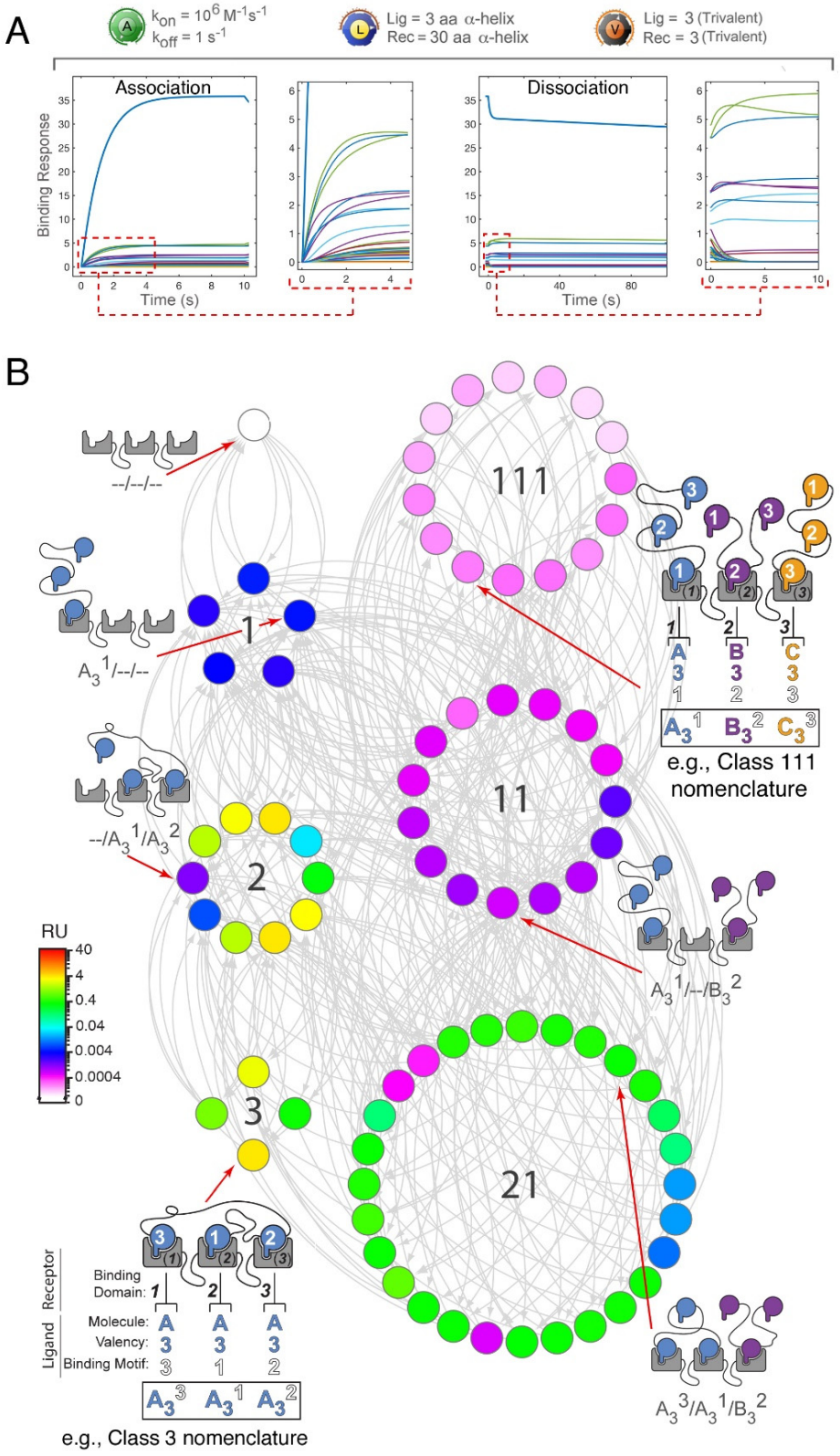
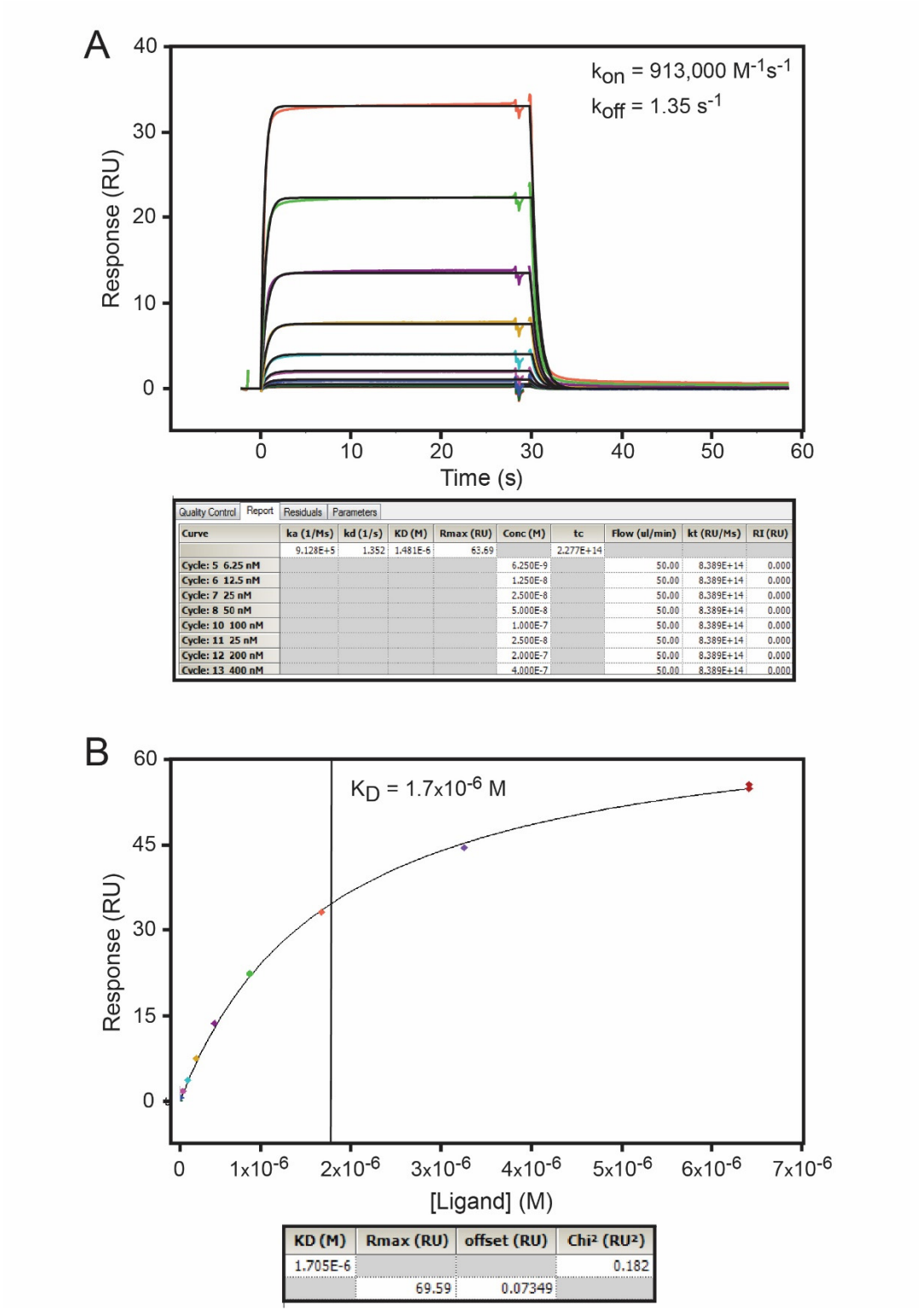


Fig. S10. The three-parameter framework for multivalency extends combinatorially to all ligand-receptor valencies. **(A)** Simulated SPR sensorgrams showing association and

dissociation kinetics. The overall response is shown in the top trace (blue), with the constituent microstates also shown. The red box insets magnify early microstate association and dissociation dynamics. **(B)** Network map of all 78 states, grouped by interaction class and color-coded by concentration at equilibrium. Depictions of representative configurations are shown for each of the configurational classes. Our derived nomenclature is shown for representative configurations in the “111” and “3” classes.



Quality Control | Report | Residuals | Parameters

Curve	ka (1/Ms)	kd (1/s)	KD (M)	Rmax (RU)	Conc (M)	tc	Flow (ul/min)	kt (RU/Ms)	RI (RU)
	9.128E+5	1.352	1.481E-6	63.69		2.277E+14			
Cycle: 5 6.25 nM					6.250E-9		50.00	8.389E+14	0.000
Cycle: 6 12.5 nM					1.250E-8		50.00	8.389E+14	0.000
Cycle: 7 25 nM					2.500E-8		50.00	8.389E+14	0.000
Cycle: 8 50 nM					5.000E-8		50.00	8.389E+14	0.000
Cycle: 10 100 nM					1.000E-7		50.00	8.389E+14	0.000
Cycle: 11 25 nM					2.500E-8		50.00	8.389E+14	0.000
Cycle: 12 200 nM					2.000E-7		50.00	8.389E+14	0.000
Cycle: 13 400 nM					4.000E-7		50.00	8.389E+14	0.000

KD (M)	Rmax (RU)	offset (RU)	Chi² (RU²)
1.705E-6			0.182
	69.59	0.07349	

Fig. S11. Surface plasmon resonance experiments performed on the monovalent SH3 - SLP76 receptor-ligand pair to obtain k_{on} and k_{off} input parameters for use in simulations. Both **(A)** kinetic and **(B)** equilibrium fits were performed.

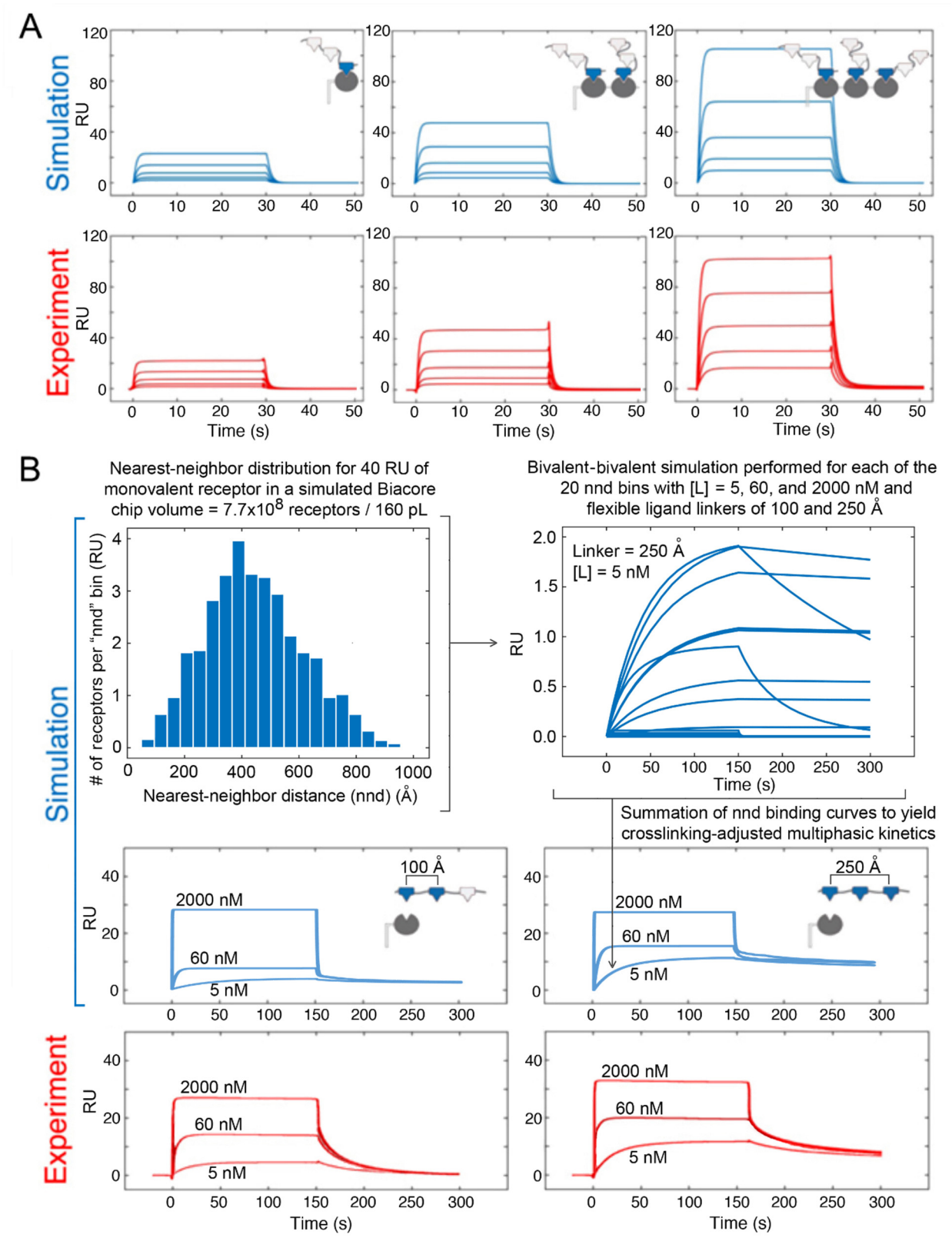


Fig. S12. Simulations of multivalent interactions based on detailed description of the configurational network. **(A)** Simulations explicitly allow for multi-ligand engagement with

a multivalent receptor. Simulated sensorgrams (blue) for monovalent ligand(s) engaging with a monovalent, bivalent, and trivalent receptor. Experiment sensorgrams (red) using monovalent, bivalent, and trivalent SH3 domain receptors with monovalent substrate peptide. **(B)** High local concentration of receptors result in multivalent-like, multi-phasic binding phenomena due to multivalent ligands both re-binding to a receptor following dissociation and crosslinking between two or more. Explicit consideration of receptor density in the simulations can account for this behavior. The immobilized receptor volume was modeled as a distribution of monovalent nearest-neighbors capable of acting as bivalent receptors. This binned distribution of receptors yields a series of binding kinetics ranging from monovalent to strongly bivalent, depending on the receptor RU and ligand linker length. Summation of this series yields a crosslinking-adjusted kinetic profile. Kinetic traces are shown for ligand concentrations of 2000, 60, and 5 nM and ligand linker contour lengths of 100 and 250 Å. Experimentally determined monovalent kinetic parameters (k_{on} = 913,000 M^{-1} s^{-1} and k_{off} = 1.35 s^{-1}) were used for the simulations.

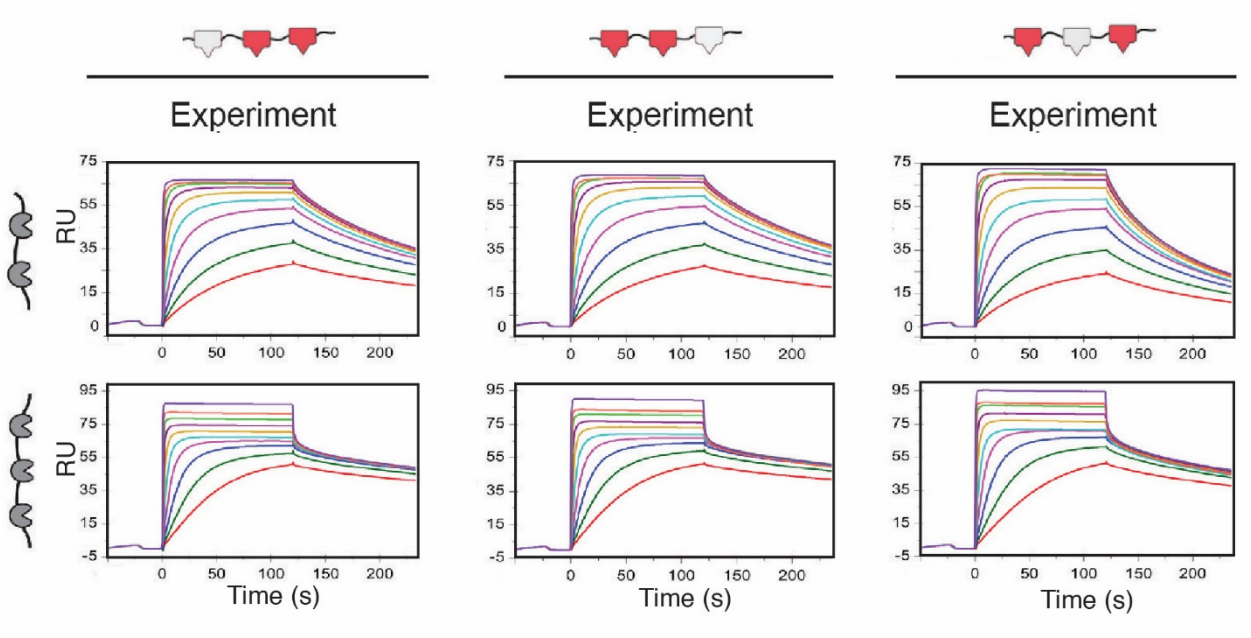


Fig. S13. The experimental system displays both subtle and significant effects on the kinetics and steady state in response to changes in topology. **(A)** SPR experiment for flexible-linker bivalent ligand against a bivalent (top panel) and trivalent (bottom panel) flexible-linker receptor. **(B)** SPR experiment as in (A) except the bivalency of ligand has been shifted N-terminal. **(C)** SPR experiment as in (A) except bivalency has been separated, creating double-length flexible linkage. Here, we see the effects of linker length on the experimental system. With a longer linker, and thus a lower [L_{eff}], a faster rate of dissociation is observed in the bivalent-bivalent interaction (panel C, top vs. panels A and B, top). Additionally, in the bivalent-trivalent interaction, longer linker lengths result in a larger population of high-stoichiometric states, indicated by the larger response signal (panel C, bottom vs. panels A and B, bottom).

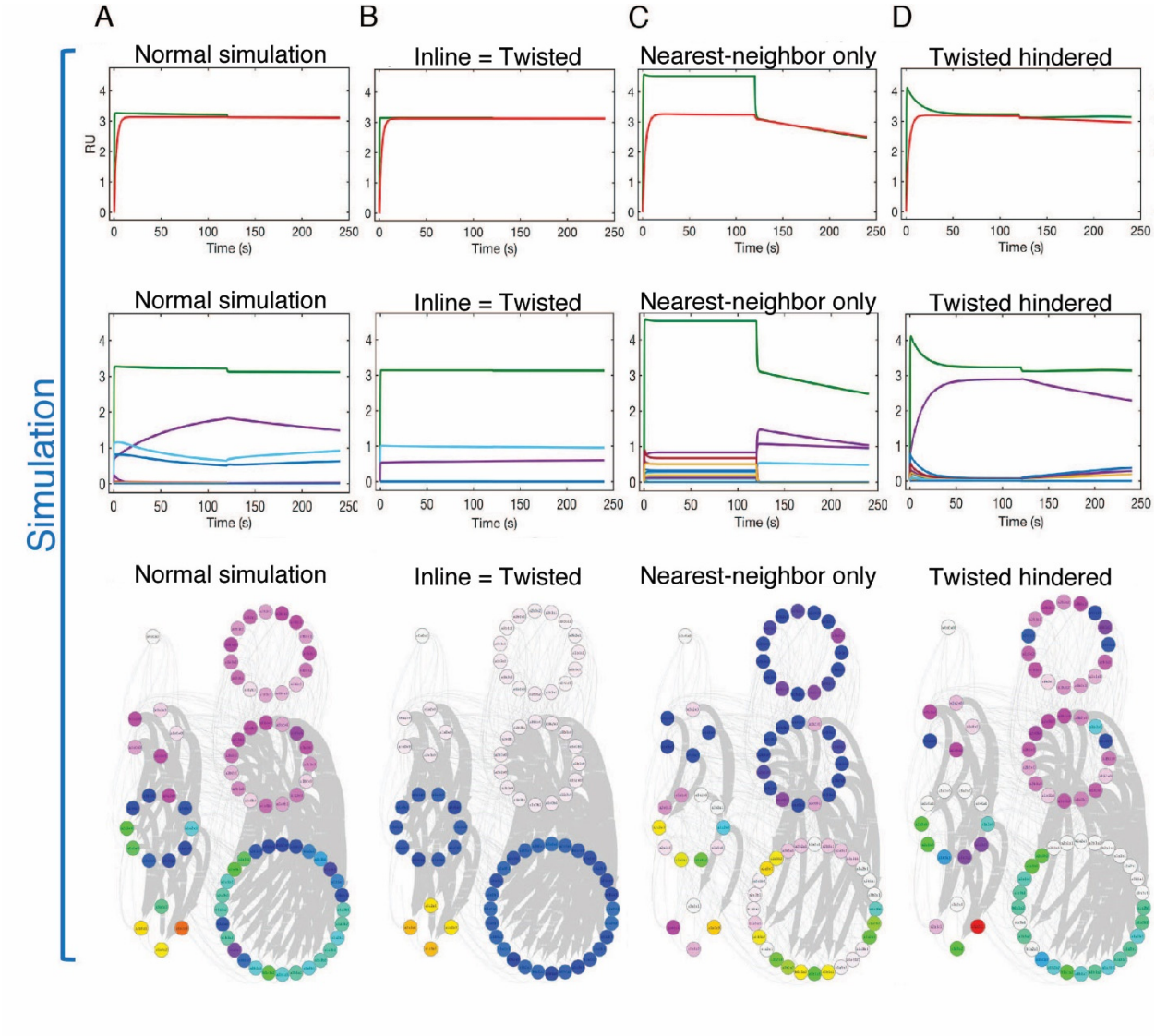


Fig. S14. Agreement between zero-fit model and experiment requires accurate description of the configurational ensemble. **(A)** Trivalent receptor-trivalent ligand simulation performed with the standard model framework described in Fig. S1. Top panel: kinetic simulations performed with 50 nM (red) and 1000 nM ligand (green). Middle panel: kinetic simulation of configurational binding states with 1000 nM ligand (green). Configurations shown are the inline trivalent configuration ($A_3^1/A_3^2/A_3^3$, purple trace) and the twisted trivalent configurations ($A_3^2/A_3^3/A_3^1$, cyan; $A_3^1/A_3^3/A_3^2$, blue). Bottom panel: network configuration at equilibrium. Permissible intracomplex configurational transitions are indicated with grey arrows. **(B)** As in (A) except network model was altered to remove steric constraints on twisted configurational transitions. **(C)** As in (A) except the multivalent network was additionally restricted by only permitting "nearest-neighbor" configuration transitions (e.g. A_3^1/A_3^2-- to $A_3^1/A_3^2/A_3^3$). **(D)** As in (A) except the network model was altered to forbid twisted configurational transitions.

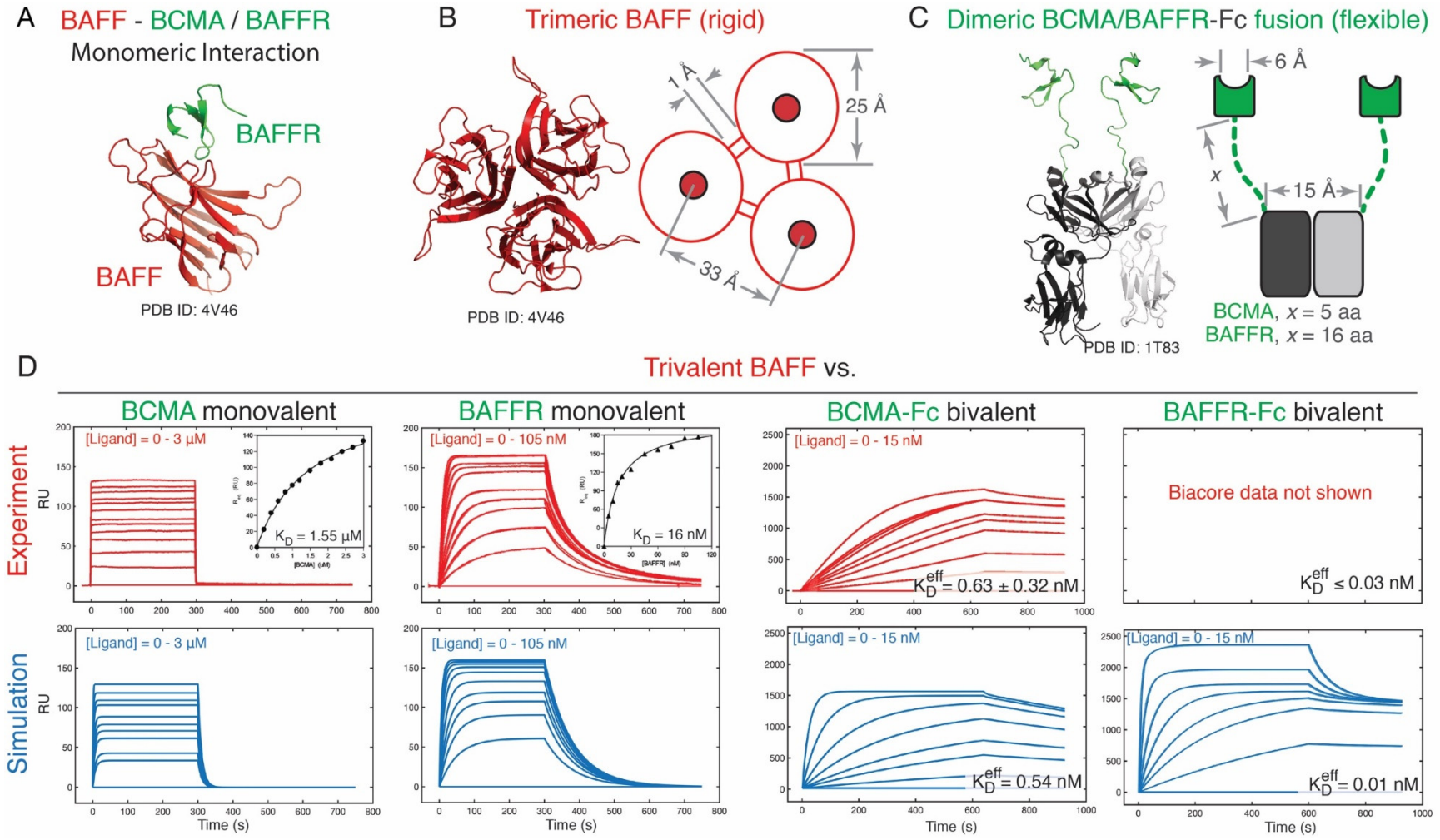


Fig. S15. Applications of the model to the multivalent experimental literature. Case Study 1: Day ES et al. Selectivity of BAFF/BLyS and APRIL for Binding to the TNF Family Receptors. *Biochemistry*, 44:1919-1931 (2005). (A) X-ray crystal structure of the BAFF-BAFFR complex (7) studied in this report. (B) The trimeric BAFF structure was used to parameterize the simulations: 25 Å rigid binding domains connected by 1 Å rigid "linkers" with a three-fold symmetry established by using one PDF calculation for the three pairwise inline distances, and one for the three pairwise twisted distances. (C) A structural

model was made for the dimeric BCMA/BAFFR-Fc dimers (8). Here, to achieve an approximation of this topology, the linker spanning the two binding domains (colored green) was given a “linker-rod-linker” construction. The “linker” represents the length “x” (taken to be flexible based on the absence of electron density in all available crystal structures); the “rod” represents the rigid 15 Å between the C-termini of the two Fc chains (colored light and dark grey). (D) Comparison of literature Biacore sensorgrams with simulations for the indicated multivalent receptor-ligand pairs. Experimental data figures adapted from Day et al. (9).

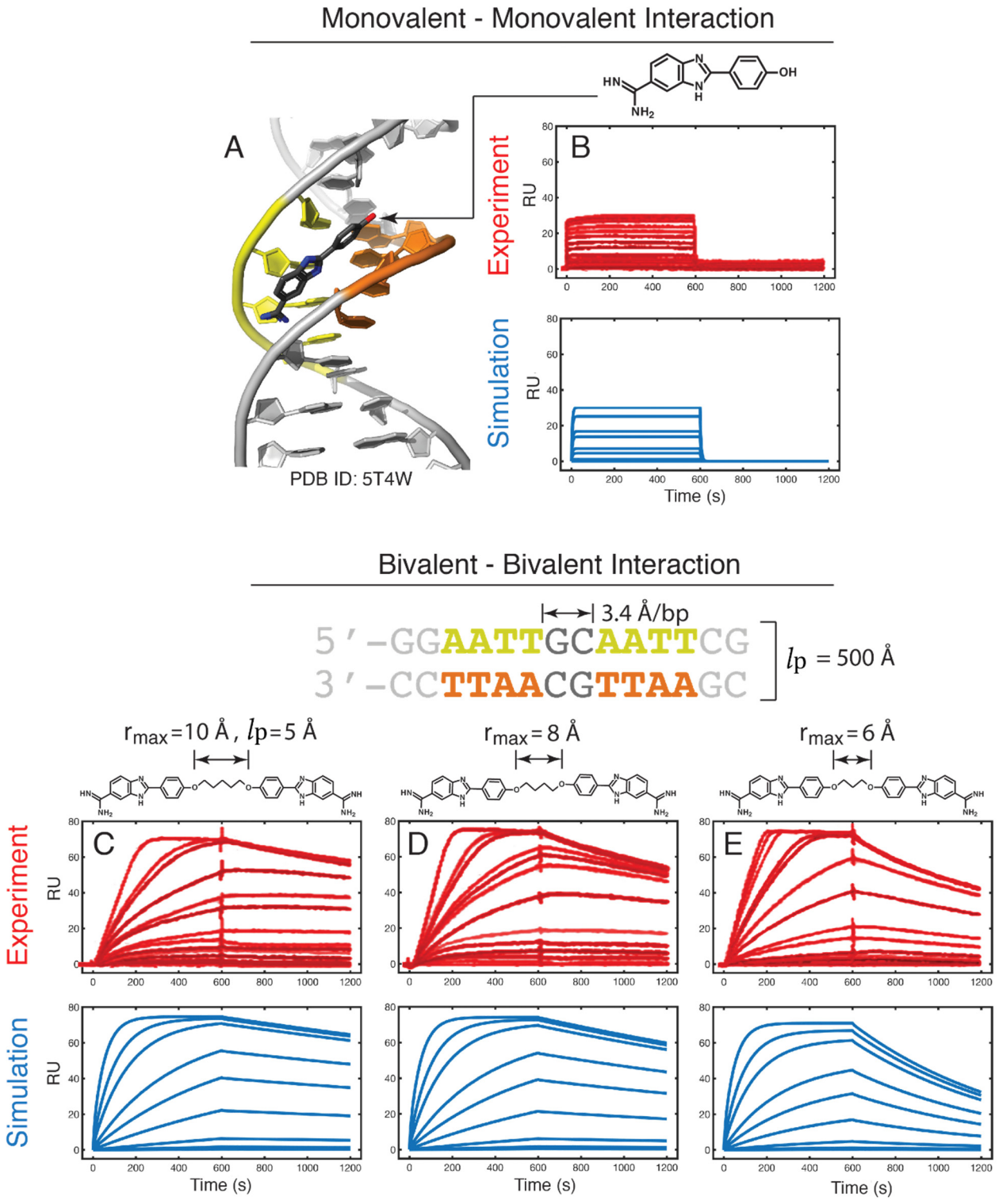


Fig. S16. Applications of the model to the multivalent experimental literature. Case Study 2: Liu et al. Designed Compounds for Recognition of 10 Base Pairs of DNA with Two AT Binding Sites. *JACS*, 134:5290-5299 (2012). (A) X-ray crystal structure of dsDNA AATT motif with an intercalated amidine-benzimidazole-phenyl (ABP) compound. (B) Experimental monovalent kinetics were compared with simulations using k_{on} and k_{off}

values of $2x10^6$ M^{-1} s^{-1} and 1 s^{-1} (10). (C-E) Bivalent $(AATT)_2$ dsDNA and $(ABP)_2$ binding experiments and corresponding simulations. The tandem AATT motifs were separated by a 2-bp GC “linker” (6.8 Å). A dsDNA molecule was treated as uniformly rigid (l_p = 500 Å) (11). Three bivalent ABP compounds were simulated with alkyl ether linkers with contour lengths of ~10, 8, and 6 Å that were modeled as highly flexible polymers (l_p= 5 Å) (12). The experimental data panels and ABP structures were adapted from Liu et al. (13).

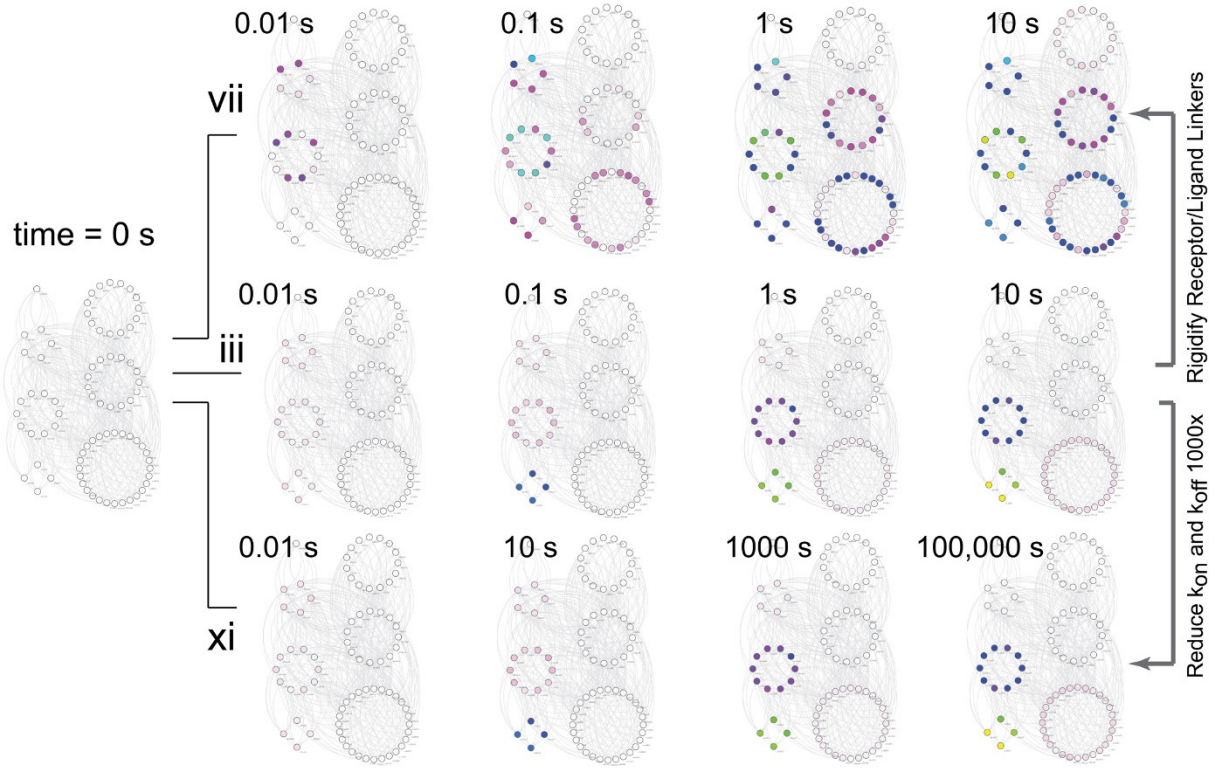
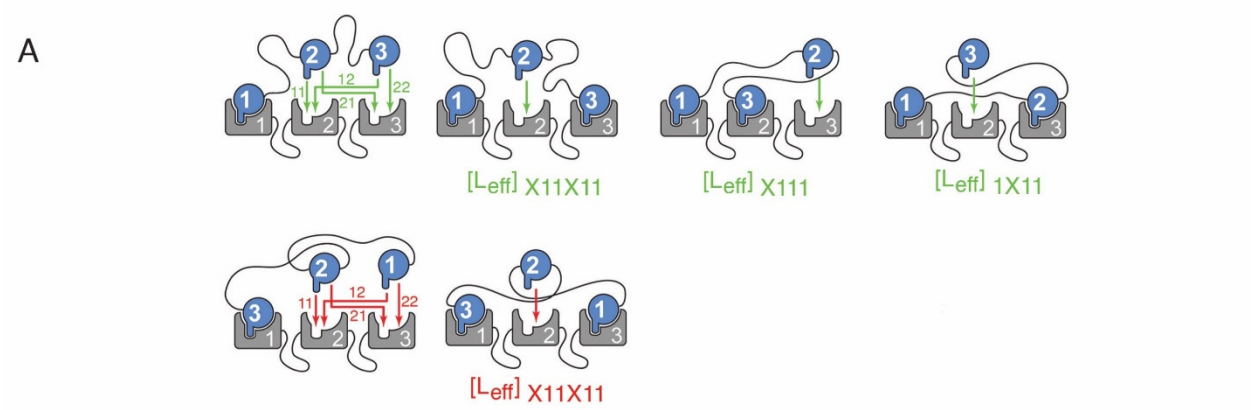


Fig. S17. Temporal evolution of network maps of vii, iii, and xi from Fig. 4 at indicated time points.



B

Effective ligand concentrations [L_{eff}] for each type of multivalent association

		i		iii		v		vii	
		Linker-driven PDF (mM)	Uniform model (mM)	Linker-driven PDF (mM)	Uniform model (mM)	Linker-driven PDF (mM)	Uniform model (mM)	Linker-driven PDF (mM)	Uniform model (mM)
Inline	$L_{eff,\ 11}$	7.4×10^{-3}	1.5×10^{-3}	1.2×10^{-2}	1.5×10^{-1}	1.1×10^{-2}	1.0×10^{-2}	2.0×10^{-7}	2.6×10^{-1}
	$L_{eff,\ 12}$	4.1×10^{-3}	1.5×10^{-3}	5.8×10^{-3}	1.5×10^{-1}	1.6×10^{-3}	1.0×10^{-2}	5.3×10^{-3}	2.6×10^{-1}
	$L_{eff,\ 1X11}$	7.4×10^{-3}	1.5×10^{-3}	1.2×10^{-2}	1.5×10^{-1}	1.1×10^{-2}	1.0×10^{-2}	2.0×10^{-7}	2.6×10^{-1}
	$L_{eff,\ 21}$	4.2×10^{-3}	1.5×10^{-3}	1.1×10^{-2}	1.5×10^{-3}	1.6×10^{-3}	1.0×10^{-2}	1.3×10^{-5}	1.0×10^{-2}
	$L_{eff,\ 22}$	2.8×10^{-3}	1.9×10^{-4}	5.2×10^{-3}	1.6×10^{-2}	1.6×10^{-3}	1.2×10^{-3}	4.5×10^{-3}	2.9×10^{-3}
	$L_{eff,\ X111}$	7.4×10^{-3}	1.5×10^{-3}	1.2×10^{-2}	1.5×10^{-1}	1.1×10^{-2}	1.1×10^{-2}	2.0×10^{-7}	2.6×10^{-1}
	$L_{eff,\ X11X11}$	7.1×10^{-3}	1.5×10^{-3}	1.1×10^{-2}	1.5×10^{-1}	1.1×10^{-2}	1.0×10^{-2}	1.8×10^{-7}	2.6×10^{-1}
Twisted	$L_{eff,\ 11}$	7.0×10^{-3}	1.5×10^{-3}	1.2×10^{-2}	1.5×10^{-1}	5.6×10^{-3}	1.0×10^{-2}	2.4×10^{-7}	2.6×10^{-1}
	$L_{eff,\ 12}$	3.9×10^{-3}	1.5×10^{-3}	5.5×10^{-3}	1.5×10^{-1}	1.6×10^{-3}	1.0×10^{-2}	4.8×10^{-3}	2.6×10^{-1}
	$L_{eff,\ 21}$	4.0×10^{-3}	1.5×10^{-3}	1.0×10^{-2}	1.5×10^{-3}	1.6×10^{-3}	1.0×10^{-2}	3.7×10^{-4}	1.0×10^{-2}
	$L_{eff,\ 22}$	2.7×10^{-3}	1.9×10^{-4}	5.0×10^{-3}	1.6×10^{-2}	1.5×10^{-3}	1.2×10^{-3}	4.1×10^{-3}	2.9×10^{-3}
	$L_{eff,\ X11X11}$	2.3×10^{-3}	1.5×10^{-3}	2.9×10^{-3}	1.5×10^{-1}	1.8×10^{-3}	1.0×10^{-2}	6.1×10^{-8}	2.6×10^{-1}

Fig. S18. Effective ligand concentrations, [L_{eff}], determined with a uniform model in which ligand is uniformly distributed in a volume with a radius of the linker contour length fails to describe the positional steric effects of both rigid and out-of-register receptor-ligand linkers that are reflected in the linker-driven PDF used in our zero-fit model. **(A)** The configurational ensemble for a trivalent receptor-trivalent ligand interaction evolves via 12 types of intracomplex association, each specified with an individual [L_{eff}] that determines the first-order rate constant of association. **(B)** Comparison of the linker-driven PDF calculation used in our model with a non-structured, uniform concentration calculation of [L_{eff}]. [L_{eff}] calculations are shown for both models using four of the simulated trivalent interactions shown in Fig. 4. Linkages within the trivalent receptor and ligand are (i) flexible, in-register; (iii) flexible, out-of-register; (v) rigid, in-register; or (vii) rigid, out-of-register.

Movie S1. Network evolution during the initial 50 seconds of the simulated association phases for trivalent receptor-ligand interactions in which twisted configurations are either allowed or hindered through alterations to the [L_{eff}] values that drive these binding conformations (see also Fig. 3).