

Binding to Nanopatterned Antigens is Dominated by the Spatial Tolerance of Antibodies

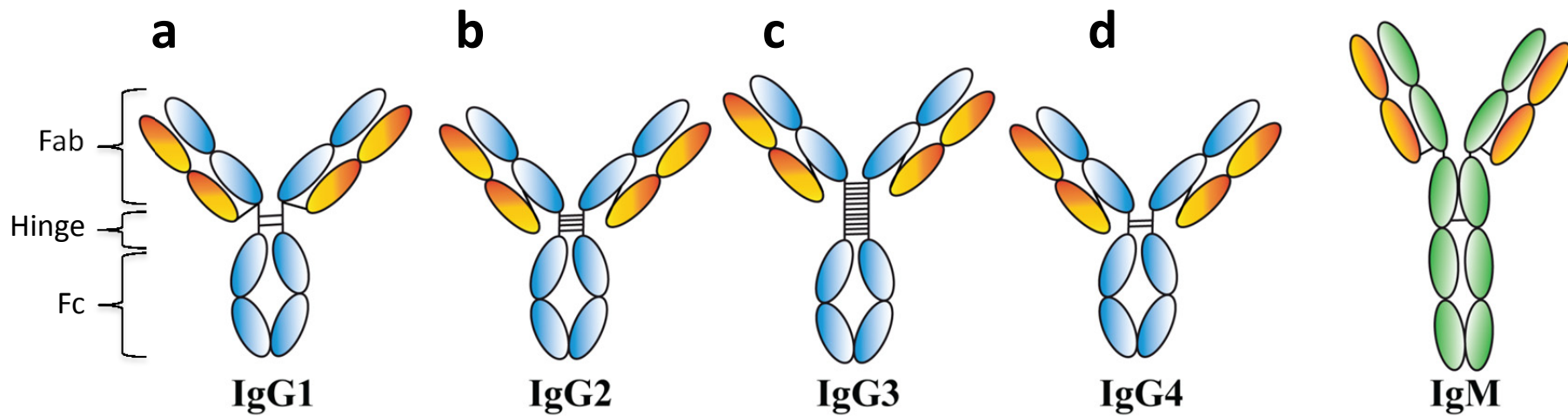
Alan Shaw, Ian T Hoffecker, Ioanna Smyrlaki, Joao Rosa, Algridas Grevys, Diane Bratlie, Inger Sandlie, Terje Enar Michaelsen, Jan Terje Andersen and Björn Högberg

Supplementary information

Item	Description	Page
Suppl. Note 1	Further on the importance of patterns in immunity	2
Suppl. Figure 1	Structure schematics of IgG1-4 and IgM	3
Suppl. Figure 2	Amino acid sequence alignment of human IgG1-4	4
Suppl. Figure 3	caDNAo design diagrams of DNA nanostructures	5
Suppl. Figure 4	Additional TEM and gel data of DNA nanostructures	6
Suppl. Figure 5	Dot plot of anti-NIP Abs data from Fig2a.	7
Suppl. Figure 6	Structural details of antigen sites on nanostructures	8
Suppl. Figure 7	Typical SPR sensorgram for nanostructure immobilization	9
Suppl. Figure 8	Typical SPR sensorgram and fitting for Ab binding	9
Suppl. Figure 9	Amino acid sequence of constant domain of rabbit IgG	10
Suppl. Figure 10	Amino acid sequence alignment of mouse IgG1 and 2a	11
Suppl. Figure 11	Rabbit IgG binding to digoxigenin patterns	12
Suppl. Figure 12	Addition of long linkers	13
Suppl. Figure 13	Significance of linkers	14
Suppl. Figure 14	Influence of nanostructure density	15
Suppl. Figure 15	Gel retardation assay for oligo incorporation estimation	16
Suppl. Figure 16	Human IgG1-4 dissociation constants vs. Separation.	17
Suppl. Figure 17	Sequence alignment of IgG3 engineered mutants	18
Suppl. Figure 18	Human monomeric IgM dissociation constants vs. separation	19
Suppl. Note 2	Details on the CTMC computer model	20
Suppl. Figure 19	Example of network state transitions	26
Suppl. Figure 20	Visualization of the breadth first search	27
Suppl. Figure 21	Transition matrices identified for different patterns	29
Suppl. Figure 22	Visualization of network of states for the hexagon	31
Suppl. Note 2 (cont.)	Solving the Markov Chain from the state transition matrix	30
Suppl. Figure 23	CTMC-model fits for the hexagon patterns	34
Suppl. Figure 24	CTMT-model state probability trajectories	35
Suppl. Figure 25	Stacking probability trajectories into macrostates	36
Suppl. Figure 26	Stacked probabilities weighted by occupancy	37
Suppl. Figure 27	Macrostates from probabilities weighted by occupancy	38
Suppl. Figure 28	State probabilities of macrostates at 0.025 nM Ab	39
Suppl. Figure 29	State probabilities of macrostates at 0.5 nM Ab	40
Suppl. Note 3	Modeling kinetics for arbitrary patterns	41
Suppl. Figure 30	Implementing distance dependence for arbitrary patterns	42
Suppl. Figure 31	Summary of steady states for example pattern	43
Suppl. Note 4	Comment on the choice of antigen–antibody pairs	44
Suppl. Table 1	Staple oligonucleotide sequences for nanostructures used	45
Suppl. Table 2	Binding kinetic data for the digoxigenin patterns	47
Suppl. Table 3	One-way ANOVA of monovalent binding samples	48
Suppl. Table 4	Binding kinetic data for human IgGs, mutants and IgM	49
Suppl. Table 5	Binding kinetic data for the hexagonal digoxigenin patterns	50

Supplementary Note 1 - Further notes and references on the importance of patterns in immunity

More examples of pattern display currently used for immunization includes the vaccine for Human Papillomavirus (HPV)⁴⁸, the Hepatitis B Virus (HBV) vaccine⁴⁹ and the malaria vaccine being tested in large clinical trials (based on the HBV vaccine structure)⁵⁰. Artificial, rationally designed, protein nanostructures are also being proposed as potential antigen carriers for vaccine applications.⁵¹ For passive therapy (*i.e.* with a target towards the effect of delivering neutralizing antibodies), bivalent binding has been found to be crucial in some instances of neutralization for example in dengue virus⁵² and H3N2 influenza⁵³. Further, the importance of Ab flexibility was recently highlighted in the context of HIV envelope glycoprotein (Env)-specific monoclonal antibody (mAb), where the virus neutralizing potency of the mAb was enhanced by artificially increasing the flexibility between the two IgG's Fab arms.^{54,55}



Supplementary Fig.1. Schematic illustrations of the human IgG1 subclasses and IgM. Crystallographic data are available for some isotypes^{10,56-58} a. IgG1. b. IgG2. c. IgG3. d. IgG4. e. Monomer of IgM. Each IgG subclass is built up of two heavy chains and two light chains that are colored blue and orange, respectively. The heavy chains and light chains are covalently paired via disulfide bridges, and so are the heavy chains via disulfide bridges of the hinge region. The four chains pair up and can be divided onto three main parts, the constant Fc that is linked to two Fab arms via the hinge region. The Fab arms harbor the specificity for a cognate antigen. The number of disulfide bridges varies between the IgG subclasses.⁵⁹ IgM exists in soluble form as hexamers and pentamers that are formed via covalent disulfide bridges between the monomers. Used in this study is a soluble monomeric version of IgM (like the one found membrane bound in B cell receptors) with its two heavy chains and light chains that are colored green and orange, respectively. IgM does not contain a classic hinge region but instead has a short stretch of amino acids with a disulfide bridge in addition to an extra constant Fc domain.⁶⁰

```

IgG1_WT   ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHFFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKE-----
IgG2_WT   ASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHFFPAVLQSSGLYSLSSVVTVPSSNFGTQTYTCNVDHKPSNTKVDKTVERK-----
IgG3b_WT  ASTKGPSVFPLAPCSRSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHFFPAVLQSSGLYSLSSVVTVPSSSLGTQTYTCNVNHKPSNTKVDKRVELKTPLGDTHTTHTCPRCPEPKSC
IgG4_WT   ASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHFFPAVLQSSGLYSLSSVVTVPSSSLGTQTYTCNVDHKPSNTKVDKRVESK-----
***** .*:*** .*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****

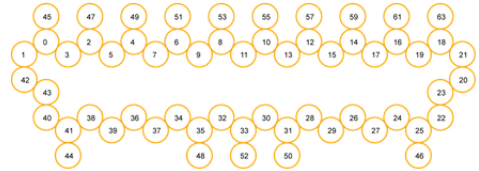
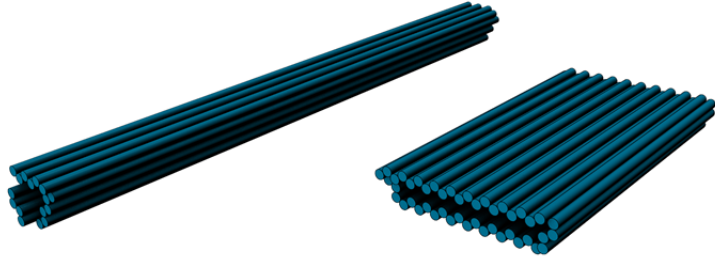
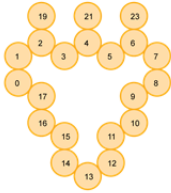
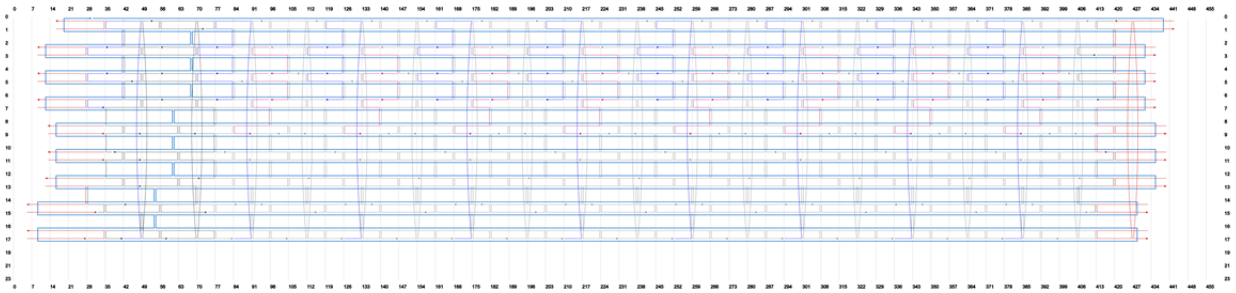
IgG1_WT   -----PKSCDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLH
IgG2_WT   -----CCVECPCPAPPV-AGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVQFNWYVDGVEVHNAKTKPREEQFNSTFRVSVLTVVH
IgG3b_WT  DTPPPCPCPEPKSCDTPPCPCRCPEPKSCDTPPCPCRCPEAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVQFNWYVDGVEVHNAKTKPREEQYNSTFRVSVLTVLH
IgG4_WT   -----YGPPCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVQFNWYVDGVEVHNAKTKPREEQFNSTYRVVSVLTVLH
***** .*:*** .*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****

IgG1_WT   QDWLNGKEYKCKVSNKALPAPIEKTIISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTFPVLDSGDSFFLYSKLTVDKSRWQQGNVFSQVMHE
IgG2_WT   QDWLNGKEYKCKVSNKGLPAPIEKTIISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTFPMLDSGDSFFLYSKLTVDKSRWQQGNVFSQVMHE
IgG3b_WT  QDWLNGKEYKCKVSNKALPAPIEKTIISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTTPMLDSGDSFFLYSKLTVDKSRWQQGNIFSCVMHE
IgG4_WT   QDWLNGKEYKCKVSNKGLPSSIEKTIISKAKGQPREPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTFPVLDSGDSFFLYSRLTVDKSRWQQEGNVFSQVMHE
***** .*:*** .*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****

IgG1_WT   ALHNHYTQKSLSLSPGK
IgG2_WT   ALHNHYTQKSLSLSPGK
IgG3b_WT  ALHNRFTQKSLSLSPGK
IgG4_WT   ALHNHYTQKSLSLSLGK
*****:***** **

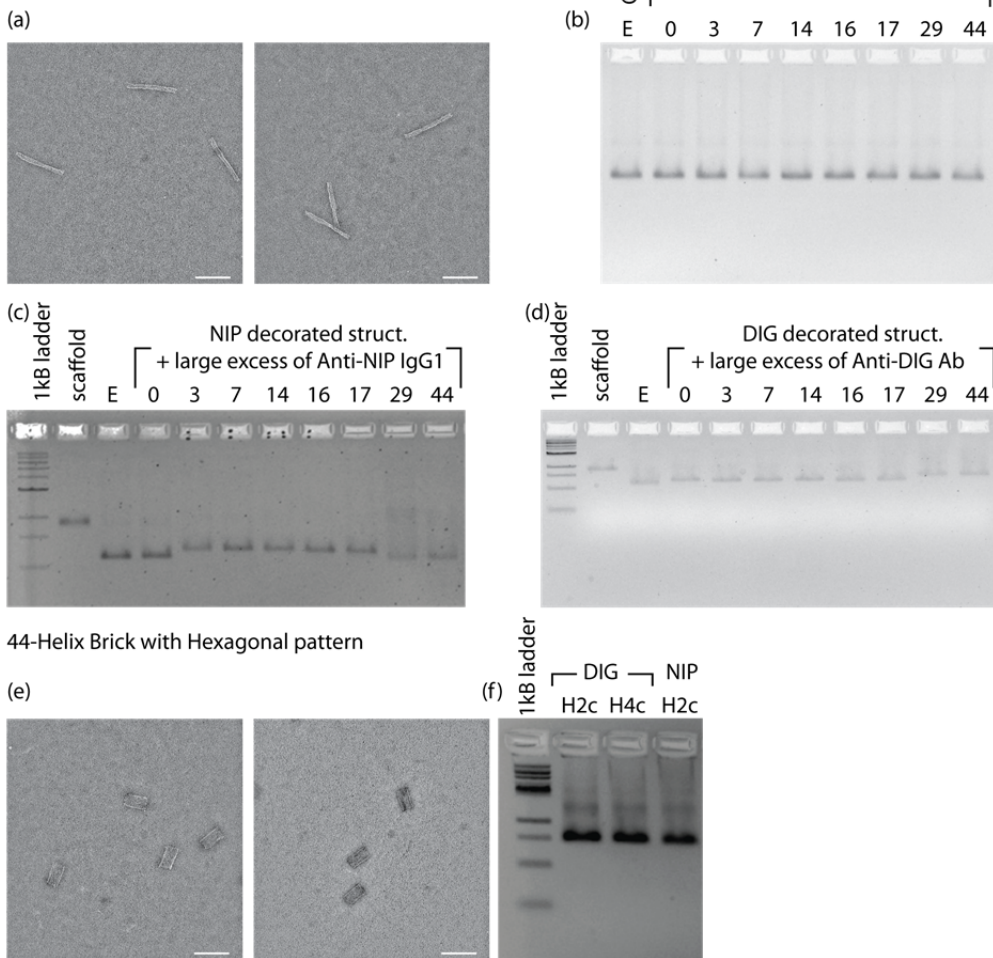
```

Supplementary Fig.2. Amino acid sequence alignment of the four human IgG subclasses. The amino acid sequences of the constant heavy chains of the four human IgG subclasses are aligned. In the current study we used the IgG1 G1m^{10,61} and IgG3 (b) allotypes. Minor differences exist in the amino acid composition of the heavy chains of the four subclasses that result in distinct differences in their ability to bind the complement factor C1q and Fc receptors. The stretch of amino acids corresponding to their hinge regions are highlighted in red. The alignment was made using the ClustalW. software.



Supplementary Fig.3. caDNA³³ design and Maya 3D rendering of the 18-helix bundle (TOP graph) and brick (Bottom graph). Each cylinder represents one double helix.

18-Helix Bundle with distance patterns



Supplementary Fig.4. Quality control of utilized structures. Transmission Electron Microscopy (TEM) (a),(e) and gel characterization (b),(c),(d) and (f) of the distance displaying 18-helix bundle and hexagon displaying brick nanostructures. All gels are 2% agarose, pre-stained with Ethidium Bromide (a) and (e): Uranyl Formate negatively stained TEM of E ('Empty' *i.e.* bearing no sites) control structures. Scale bars are 100 nm. The 18-Helix Bundle is well characterized in ref (34). The 44-Helix Brick have been folded more than ten times and imaged with TEM for three times.

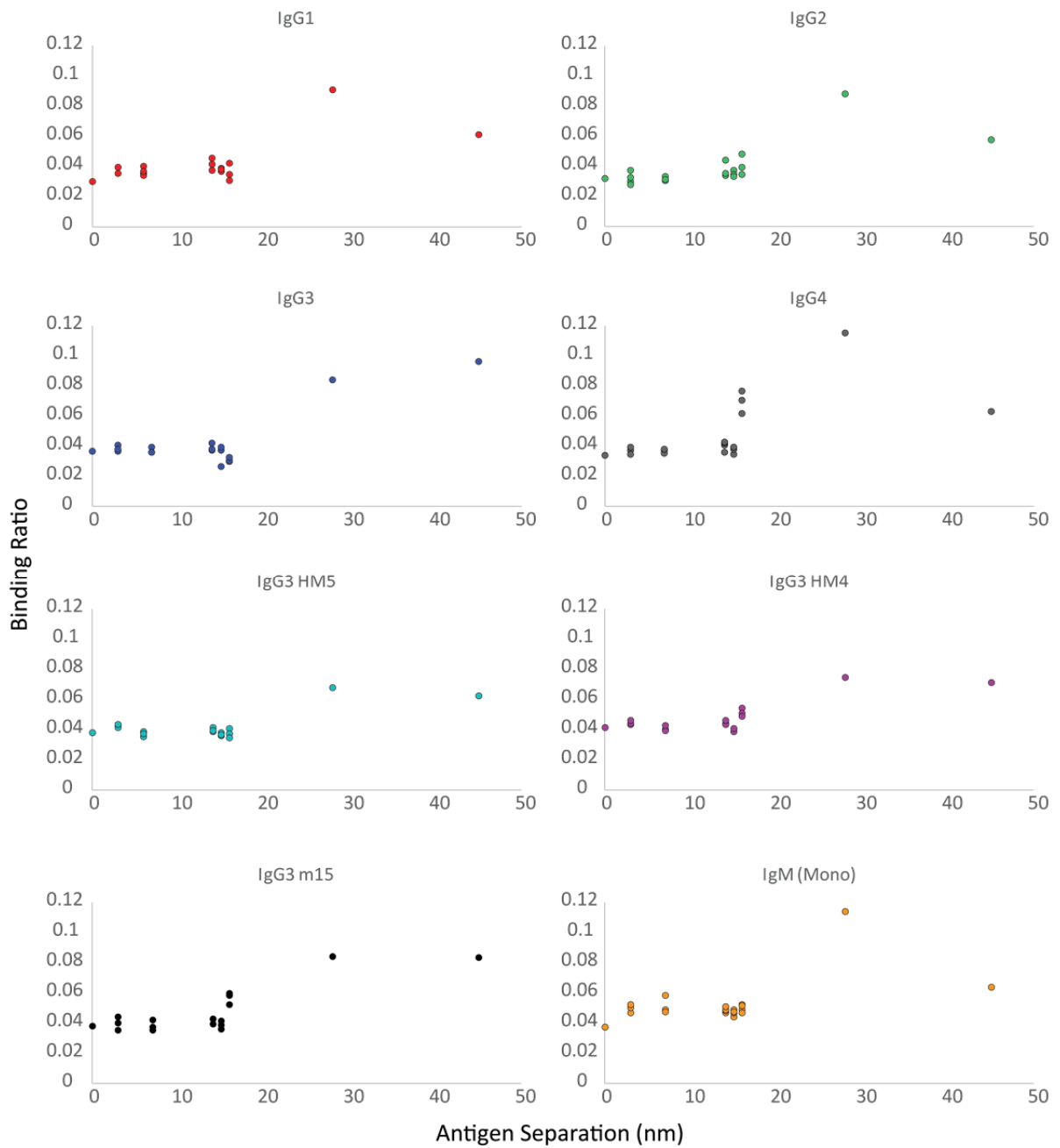
(b) Identical gel mobility of the distance displaying 18-helix bundle reveals no discernable difference between the distance cases indicating proper folding of all structures. Empty, E, structure bearing no protruding sites. Other lanes are the structures used for the NIP distance experiments.

(c) Structures used for the NIP distance experiments incubated with an excess (1:2000 structure:Abs 0-17 nm, 1:4000 29 & 44 nm) of anti-NIP human IgG1. Appreciable binding and resulting shift can be seen in the bivalent capable samples (3-17 nm) whereas the predominantly monovalent samples (0, 29 and 44 nm) display bare discernable shifts due to the low affinity at monovalent binding.

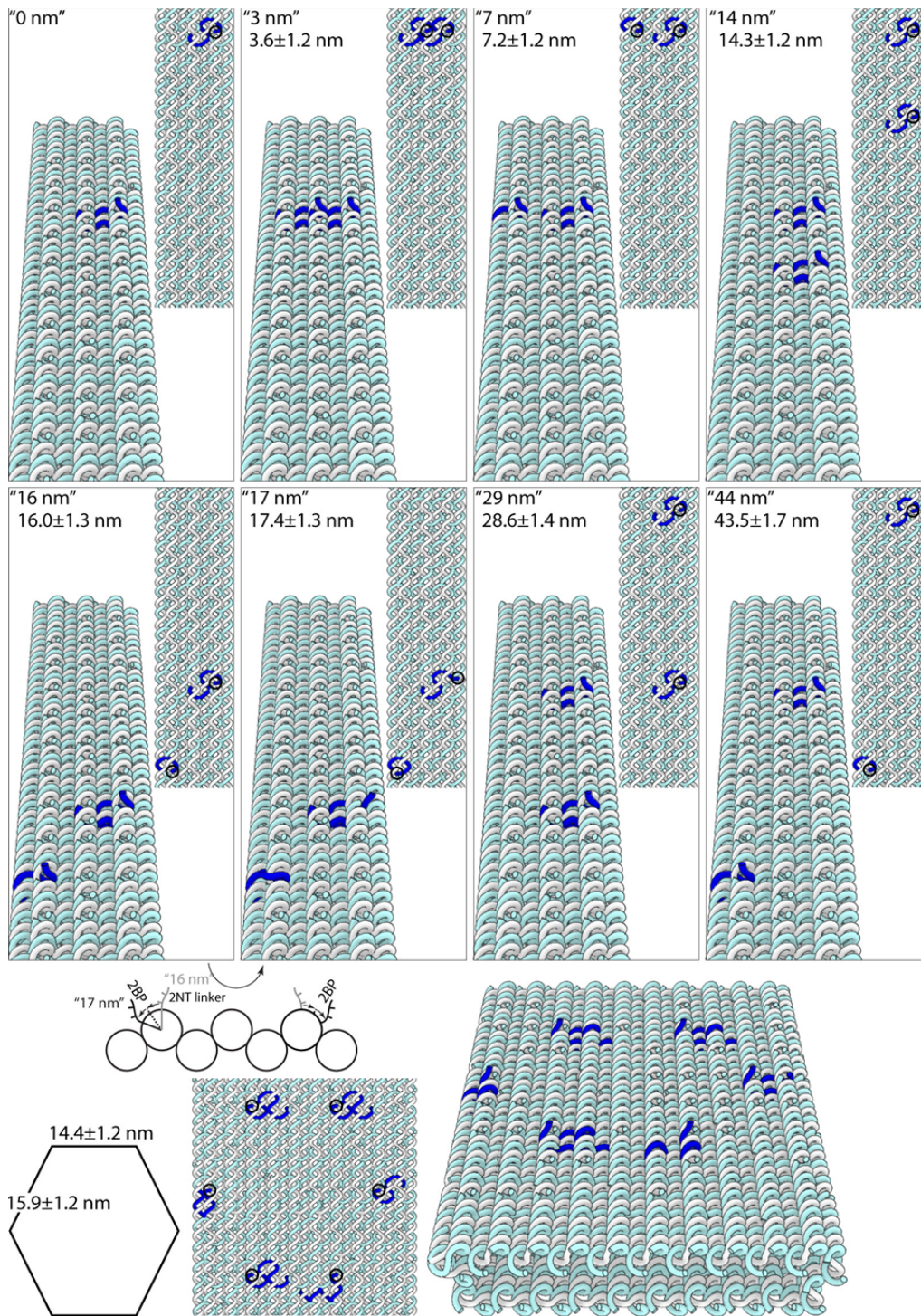
(d) Using the higher affinity rabbit anti-DIG Abs with DIG decorated structures reveals shifts also for the monovalent samples (0, 29 and 44 nm) with a larger mobility shift for the longer distances (29 and 44 nm) indicating two monovalently bound Abs. Excess is 1:1000 structure:Abs 0-17 nm, 1:2000 29 & 44 nm.

(f) Agarose gel of the hexagonal pattern displaying brick structures without Abs showing no discernable differences in folding quality.

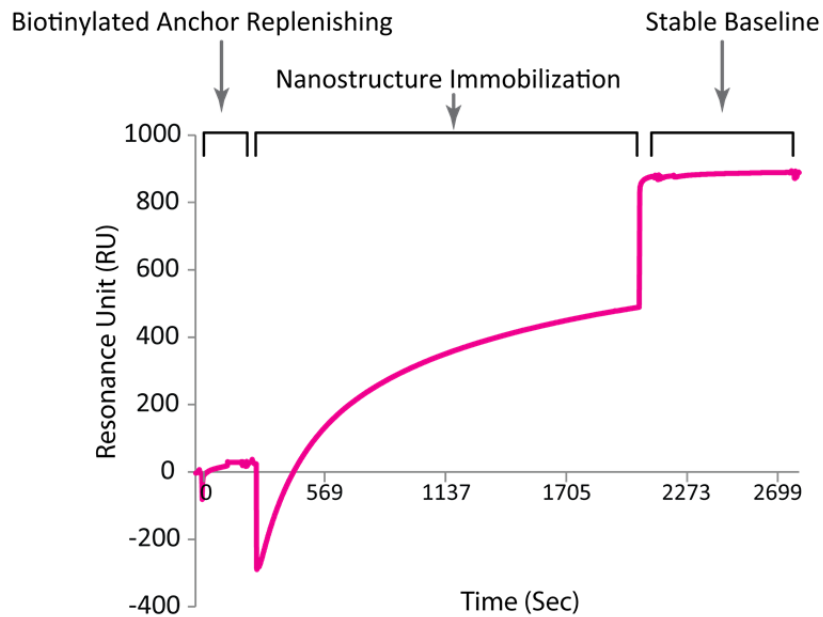
The gels in b-d are repeated three times with similar results.



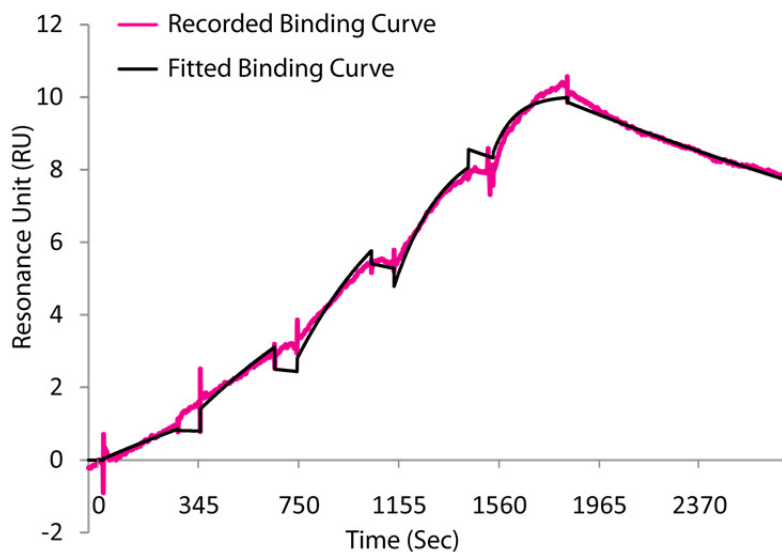
Supplementary Fig. 5 Dot plot of anti-NIP Abs data from Fig2a. Dot plot of the binding ratio column plots presented in figure 2a and b. Monovalent binding samples (0, 28, 45 nm) $n = 1$; bivalent binding samples (3, 7, 14, 16, 17 nm) $n = 3$ or 4, details provided in **Supplementary Table 4**. Each dot represents values from individual experiments.



Supplementary Fig. 6. Protruding sites presenting antigens on the 18HB and brick and distances between the antigen sites. Perspective and top views shown with potential antigen carrying staples marked in blue. The distances are estimated using TEM validated values from ¹⁸ with 0.34 nm per base-pair along the helical direction and 2.1 nm per helix perpendicular to the helical direction. Given the flexibility of the structures and the 2 nt linker that attaches the antigens in most cases, an uncertainty of at most 2 nm can be expected for all distances. This, in combination with the spatial variance found in earlier work on the 18HB³⁶ gives the indicated standard deviations. (See also Methods.) The indicative values of 0, 3, 7, 14, 16, 17, 29 and 44 nm are used as estimates of the underlying structure. The 17 nm structure is created from the 16 nm one by rotating the protruding ends by two BPs (see scheme under the 16 and 17 nm structures). The 0 nm structure contains only one antigen site. Distances shown are design distances \pm the distance error as described in methods.



Supplementary Fig. 7. Representative SPR sensorgram showing capturing of antigen decorated and nanostructures to an oligonucleotide immobilized CM3 chip surface. This sensorgram is very reproducible ($n > 200$), and nanostructure immobilization sensorgrams always look like the one as shown above.



Supplementary Fig. 8. An example of a sensorgram showing injection of serial dilutions of an Ab over a surface immobilized with nanostructures decorated with the cognate antigen. A single-cycle kinetic program was used and the resulting binding curve was fitted to the 1:1 Langmuir binding model supplied with BIAcore T200 evaluation software. This sensorgram is very reproducible ($n > 200$), and antibody binding curves (anti DIG and bivalent anti NIP) always look like the one as shown above. For weaker binding curves (Monovalent NIP and both monovalent and bivalent NP), binding reaches steady state rapidly and the binding curves look like square pulses, which corresponds to the steady state response, and when plotted against antibody concentration the K_D can be extracted by fitting the plot with a 1:1 steady state binding model.

GQPKAPSVFPLAPCCGDTSSSTVTLGCLVKGYLPEPVTVTWNSGTLTNGVRTFP SVRQSSGLYSLSSVVSVTSSSQPV
TCNVAHPATNTKVDKTVA **PSTCSKPTCP** PPELLGGPSVFI FPPKPKDTLMISRTPEVTCVVVDVSQDDPEVQFTWYIN
NEQVRTARPPLREQQFNSTIRVVSTLPITHQDWLRGKEFKCKVHNKALPAPIEKTI SKARGQPLEPKVYTMGPPREEL
SSRSVSLTCMINGFYPSDISVEWEKNGKAEDNYKTTPAVLDSG SYFLYNKLSVPTSEWQRGDVFTCSVMHEALHNHY
TQKSI SRSPGK

Supplementary Fig.9. The amino acid sequence of the heavy chain constant domains of rabbit IgG. The short stretch of amino acids corresponding to the hinge region is highlighted in red.

```

mIgG1      AKTTPPSVYPLAPGSAAQTNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLESD
mIgG2a     AKTTAPSVYPLAPVCGDTTGSSVTLGCLVKGYFPEPVTTLTWNSGSLSSGVHTFPAVLQSD
          ****.***** .. *. * *****:*****:****

mIgG1      LYTLSSSVTVPSSPRPSETVTCNVAHPASSTKVDKKIVPRDCGCKPCICTVPEVS-----
mIgG2a     LYTLSSSVTVTSSTWPSQSITCNVAHPASSTKVDKKIEPRGPTIKPCPPCKCPPAPNLLGG
          *****.*. **:::***** ** . *** ..

mIgG1      -SVFIFPPKPKDVLITITLTPKVTCVVVDISKDDPEVQFSWFVDDVEVHTAQTQPREEQFN
mIgG2a     PSVFIFPPKIKDVLMISLSPIVTCVVVDVSEDDPDVQISWFVNNVEVHTAQTQTHREDYN
          ***** **** *::* *****:*:***:***:*****:*****.:*::*

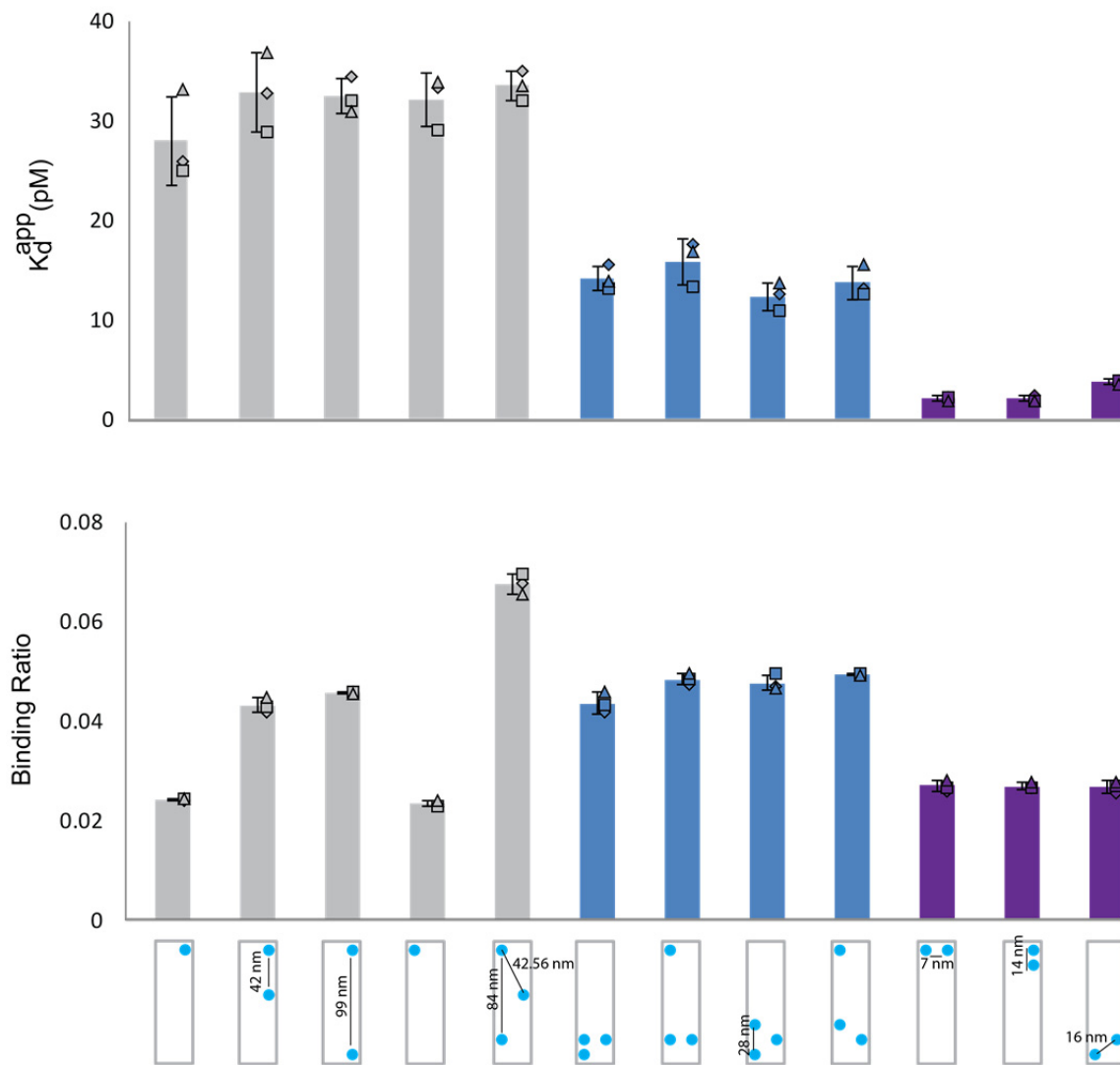
mIgG1      STFRSVSELPIMHQDWLNGKEFKCRVNSAAFPAPIEKTISKTKGRPKAPQVYTIPPPKEQ
mIgG2a     STLRVVSALPIQHQDWMSGKEFKCKVNNKDLPAPIERTISKPKGSVRAPQVYVLPPEEE
          **: * ** ** *:*.*****:*. :*****:*****.* :*****.:***:*.

mIgG1      MAKDKVSLTCMITDFFPEDITVEWQWNGQPAENYKNTQPIMNTNGSYFVYSKLNVQKSNW
mIgG2a     MTKKQVTLTCMVTDFMPEDIYVEWTNNGKTELNYKNTEPVLDSDGSYFMYSKLRVEKKNW
          *:*.**:*****:***:***** ** ***: . *****:*::::*****:*****.*:*.**

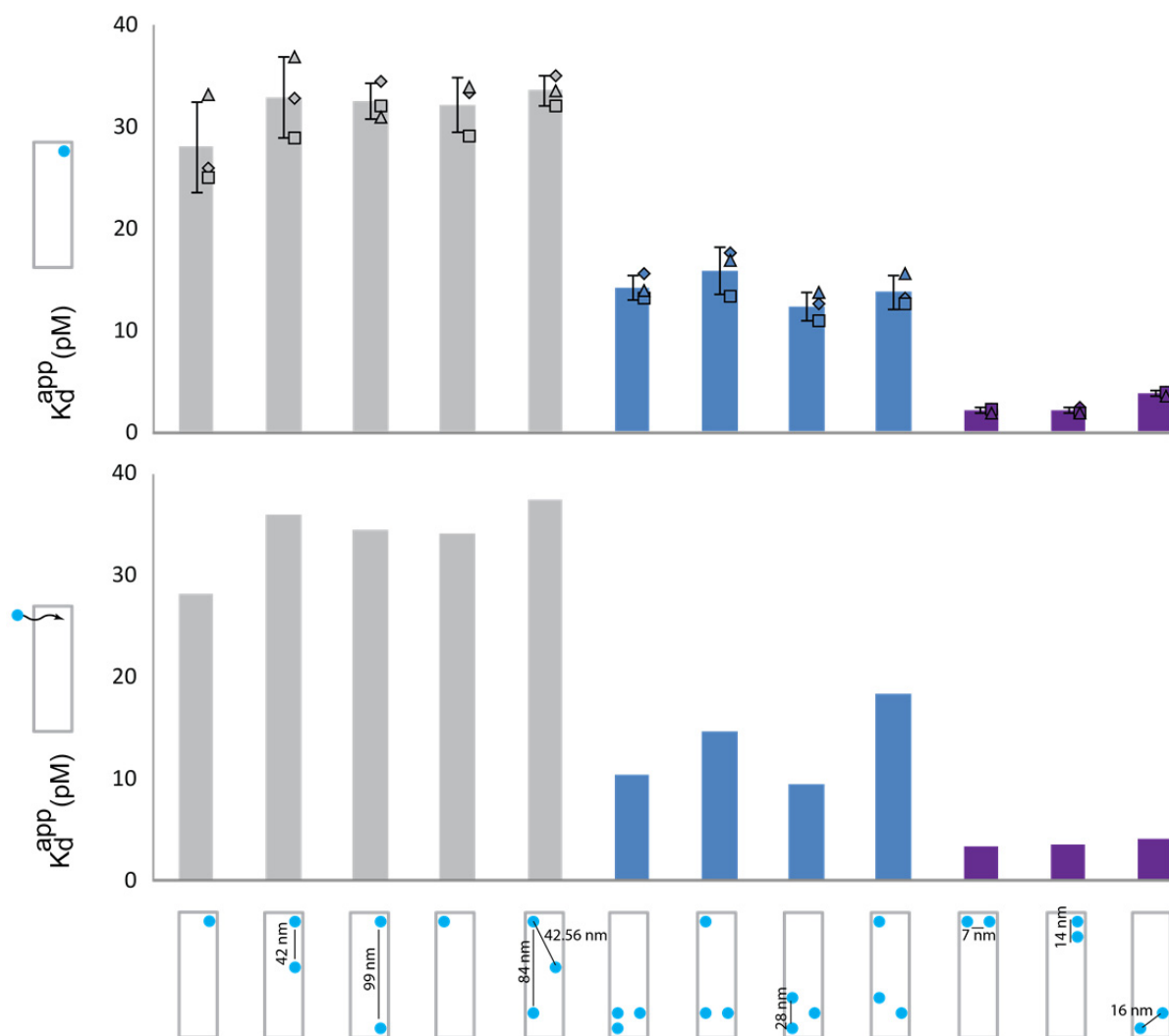
mIgG1      EAGNTFTCSVLHEGLHNHHTTEKSLSHSPGK
mIgG2a     VERNYSYSCSVVHEGLHNHHTTKSFSRTPGK
          *:::***:***** ***:*:***

```

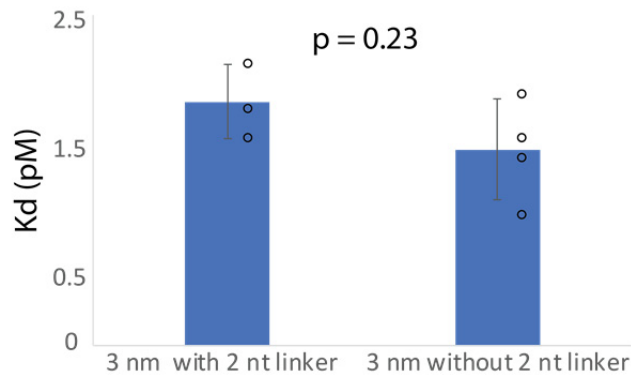
Supplementary Fig.10. The amino acid sequence of the constant domains of heavy chain of mouse IgG1 and IgG2a. The stretch of amino acids corresponding to their hinge region are highlighted in red. The alignment was made using the ClustalW. software.



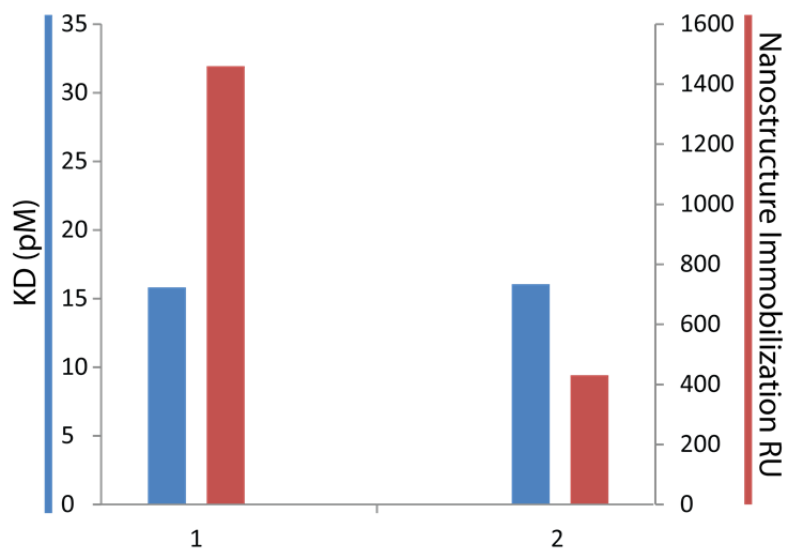
Supplementary Fig.11. Control pSPR experimental with digoxigenin nanopatterns and rabbit anti-digoxigenin IgG. Grey squares represent the 18 helix bundle while blue dots represent digoxigenin molecules. **a.** Control experiments were carried out with structures decorated with one, two or three antigens. Monovalent samples were labeled with one antigen or multiple antigens separated at distances that do not allow bivalent binding. Bivalent samples consist of two antigens at distances that allow bivalent binding; trimers are separated into two classes; three closely packed antigens and each antigen has two nearest neighbors, and a pair of separated monomer and dimer. The monovalent samples resulted in relatively weak binding affinities with binding occupancy that scales well with the number of antigens per structure. Bivalent samples showed significant increase in the binding affinity with a binding occupancy of roughly one Ab per structure, indicating successful formation of bivalent binding. The separated trimers and the closely packed trimers showed similar binding affinity, with values between weak monovalent binding and strong bivalent binding. The binding occupancy resulted in roughly two Abs per structure, indicating a 1:1 mixture of monovalent and bivalent binding. The normalized binding capacity is calculated as follows: we first calculate the binding ratio (maximum RU for Ab binding/RU on immobilized nanostructures) and then we normalize the monovalent (0 nm, 18HB decorated with a single digoxigenin) sample to 1, and all other samples are adjusted accordingly. (n= 3; columns represent the average value; error bars represent the standard deviation; dots represent data from individual repeats)



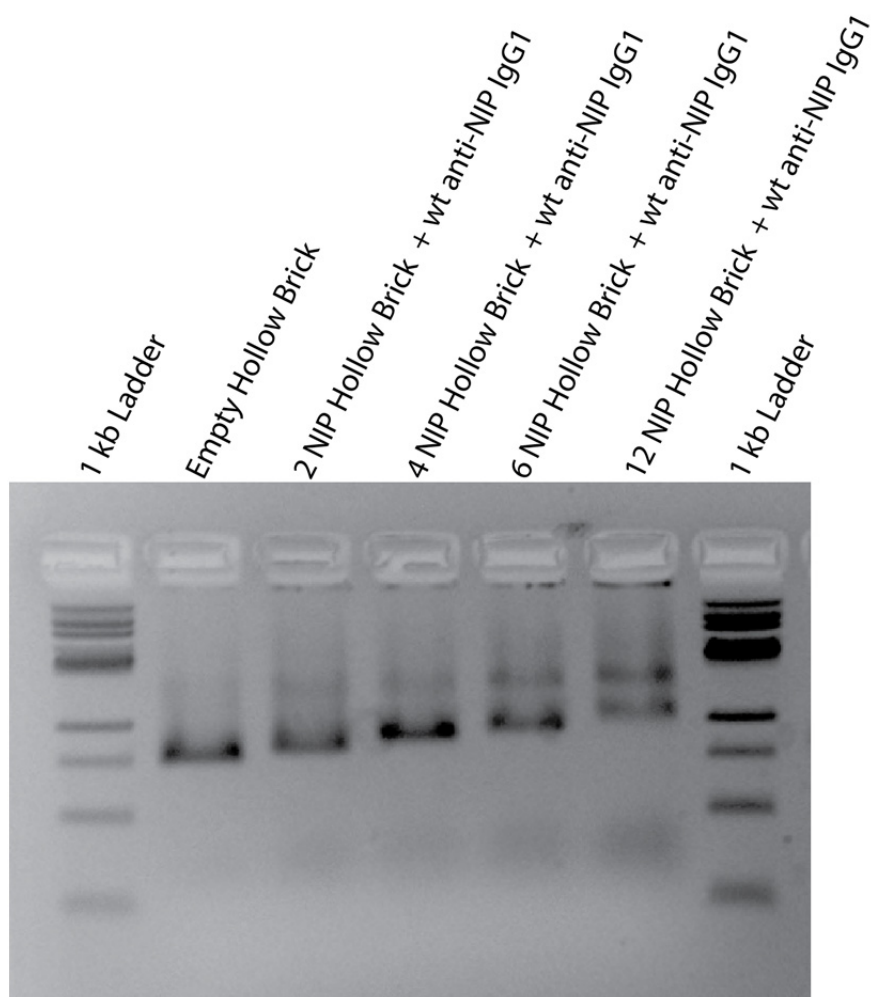
Supplementary Fig.12. Comparison between the digoxigenin nanopatterns in Supplementary Fig.7 with or without the addition of flexible linkers. The addition of flexible 21 oligonucleotide linkers to the previously tested digoxigenin nanopatterns resulted in similar results: monovalent samples resulted in weak monovalent binding; bivalent samples resulted in strong bivalent binding; antigen trimers resulted in a rough 1:1 mixture of monovalent and bivalent binding, which indicate that the binding energy can be enough to cope with the flexibility of the target. (Top panel: $n = 3$; columns represent average value; error bars represent standard deviation. Bottom panel: $n = 1$. Dots represent data from individual repeats)



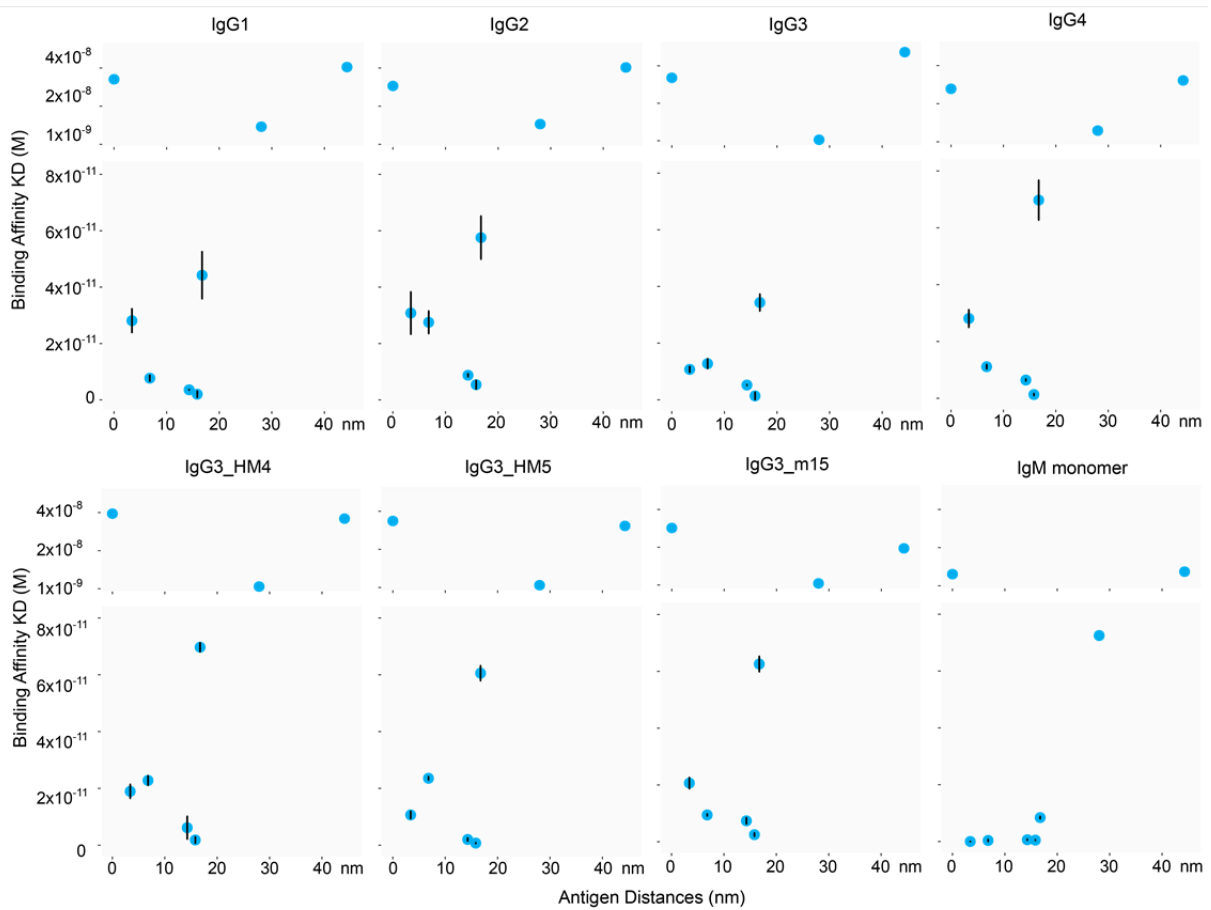
Supplementary Fig.13. Comparison between the binding affinity of rabbit anti-DIG IgG towards two DIG placed at 3.4 nm, with or without a 2 nt linker showing no significant difference between the two cases. (n = 3; column represent average; error bars represent standard deviation; two tailed equal variance student's t test was used and p = 0.23. Dots represent data from individual repeats)



Supplementary Fig.14. Comparison of the antigen binding affinity of anti-digoxigenin rabbit IgG at two different nanopattern immobilized levels. The results showed that different antigen nanopattern immobilized levels (red columns) did not affect binding affinity (blue columns). The structure used was the 18 helix bundle decorated with one digoxigenin molecule. (n = 1)



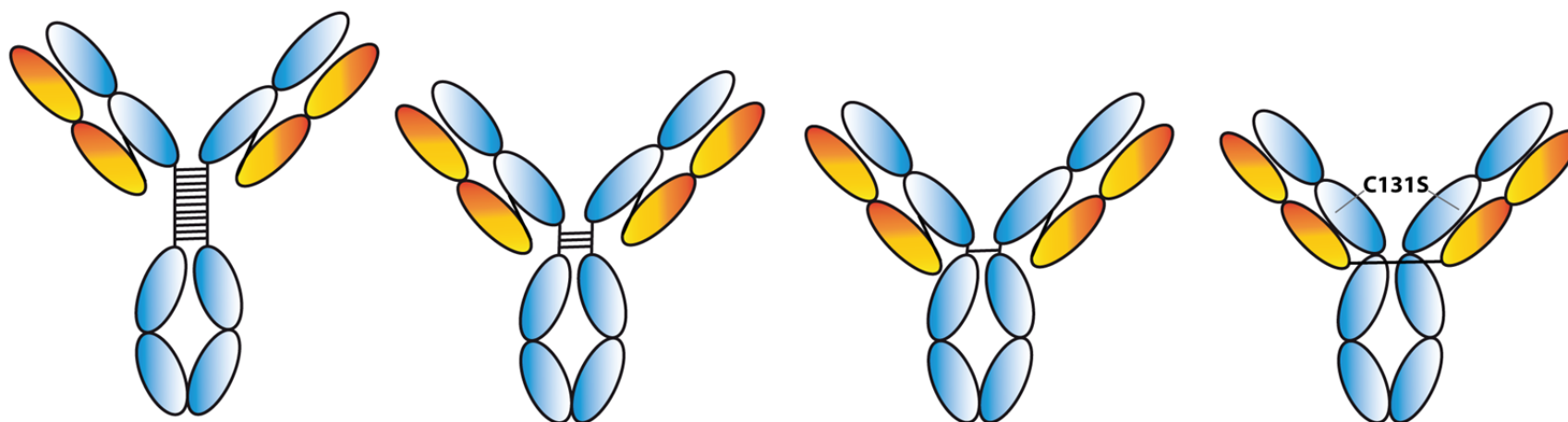
Supplementary Fig.15. Gel retardation assay showing quantitative incorporation of NIP modified oligonucleotides to the brick nanostructure. The nanostructure concentration is 5 nM and human anti-NIP IgG1 was added as 10X excess to each NIP molecule, which will result in a final antibody concentration in the range of 100-600 nM and at the concentration no monovalent binding will form. If there are defects in the NIP incorporation, for the 2 NIP structure we should see a mixture of no and one antibody, for the 4 NIP structure we should see a mixture of zero, one and two antibodies, and so forth. However, only single bands were observed, indicating that there are no significant defects in the incorporation yield of NIP modified nanostructures. (n = 1)



Supplementary Fig.16. Plotting of the NIP distance on the nanostructures against the derived binding affinity of the human anti-NIP Abs with the variations in the distances included. (n = 3; dots represent average value, vertical error bars represent standard deviation; horizontal error bars represent the flexibility of the ssDNA linkers. For dots without vertical error bars n = 1)

	<u>17</u>	<u>15</u>	<u>15</u>	<u>15</u>
IgG3	ELKTPLGDTTHT	CPRCPEPKS	CDTPPP	CPRCPEPKS
m15	EPKSCDTPPPCPRCP
HM4AACAA
HM5

(a)



(b)

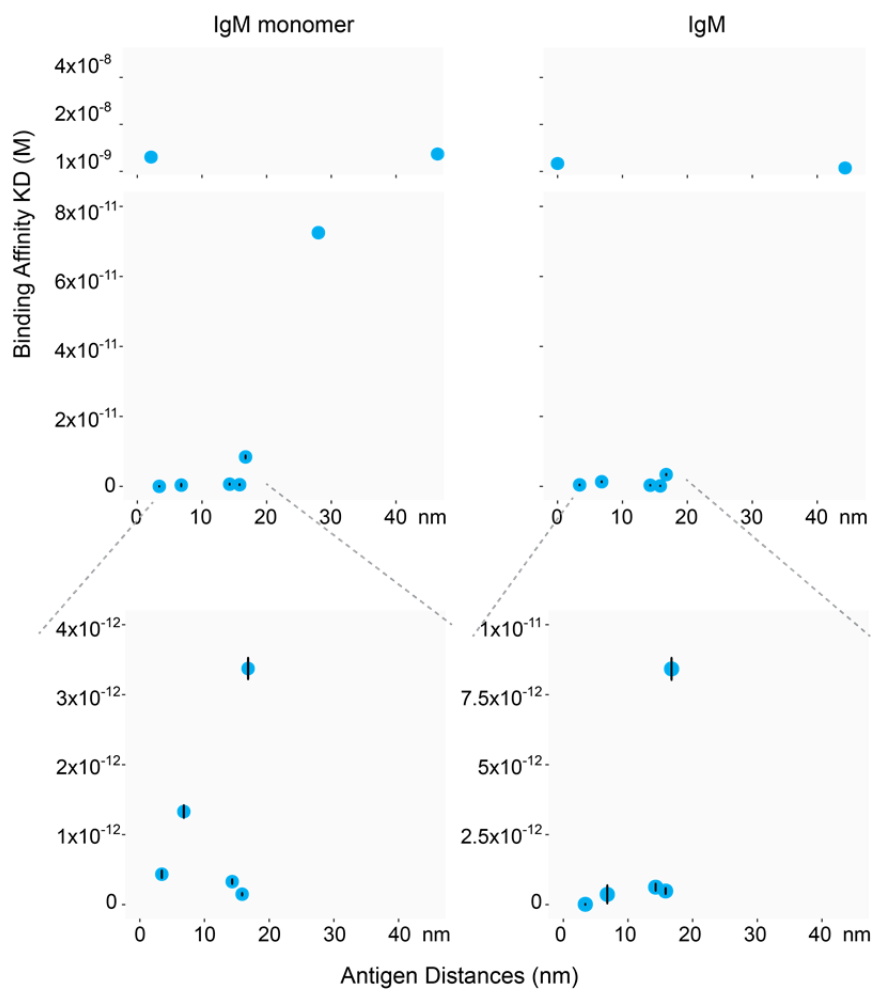
IgG3

m15

HM4

HM5

Supplementary Fig.17. a. Amino acid sequence alignment of human IgG3 and hinge-engineered variants. The amino acid sequence of the hinge region of the IgG3 (b) allotype is shown and compared with the three hinge-engineered IgG3 variants, m15, HM4 and HM5.⁴⁷ The IgG3 hinge is encoded by four exons where the first exon codes for 17 amino acids while exons 2, 3 and 4 are identical and code for 15 amino acids each. Thus, IgG3 can be described as having a 17-15-15-15 hinge region. The m15 variant only contains one of the 15 amino acid encoding exons while HM5 lacks the whole 17-15-15-15 hinge region and also lacks a disulfide bridge between their light and heavy chains due to a cysteine to serine mutation at amino acid position 131 of the constant domain 1 (C131S) (not shown). This allows the two light chains to form an inter-chain disulfide bridge through their C-terminal cysteine residues. HM5 is thus “hingeless”. For HM4, the native genetic hinge is replaced with a short but potentially flexible synthetic segment encoding the amino acids Ala-Ala-Ala-Cys-Ala. The cysteine residues of the hinge are highlighted in red. **b.** Schematic illustrations of the human IgG3, m15, HM4 and HM5.



Supplementary Fig.18. Plotting of the NIP distance on the nanostructures against the binding affinities derived from monomeric anti-NIP IgM (C575S) and pentameric IgM. ($n = 3$; dots represent average values; error bars represent standard deviation. For dots without error bars $n = 1$).

Supplementary note 2. Continuous Time Markov Chain (CTMC) Model

Deriving the units of the transition rates and SPR signal in the CTMC model

The CTMC model computes the probabilities of states connected by transition rates that have units of sec^{-1} . The following is a derivation of the relationship between measurable physical quantities and those of the model.

Consider a simple monovalently configured antigen pattern where each origami can be either empty or occupied by a single, monovalently bound antibody. The system can be represented by the following chemical formula



where $[AB]$ represents the molar concentration of antibodies, $[S_-]$ is the molar concentration of empty nanostructures, $[S_{AB}]$ is the molar concentration of bound antibody-structure complexes, and k_1 and k_{-1} are the forward and reverse rates respectively.

We can define a total structures quantity as the sum of all species concentrations that contain structures, i.e. omitting free antibodies:

$$[S_{tot}] = [S_-] + [S_{AB}] \quad \text{Eq 2}$$

The above expression can be generalized for n states, where each state necessarily contains 1 structure.

$$[S_{tot}] = [S_0] + [S_1] + \dots [S_n] \quad \text{Eq 3}$$

The probability of each species containing a structure is the fraction of that species relative to the total number of structures.

$$p_{s_i} = \frac{[S_i]}{[S_{tot}]} \quad \text{Eq 4}$$

Thus we have

$$p_{s_-} = \frac{[S_-]}{[S_{tot}]} \quad \text{Eq 5}$$

and

$$p_{s_{AB}} = \frac{[S_{AB}]}{[S_{tot}]} \quad \text{Eq 6}$$

Together these satisfy the normalization condition that all probabilities should add to 1, representing all structure-containing species, and excluding free antibodies.

$$1 = \frac{[S_-]}{[S_{tot}]} + \frac{[S_{AB}]}{[S_{tot}]} \quad \text{Eq 7}$$

and the generic form:

$$1 = \sum_{i=1}^n \frac{[S_i]}{[S_{tot}]} \quad \text{Eq 8}$$

From the reaction equation, we can write differential equations for each species

$$\frac{d[AB]}{dt} = -k_1[AB][S_-] + k_{-1}[S_{AB}] \quad \text{Eq 9}$$

$$\frac{d[S_-]}{dt} = -k_1[AB][S_-] + k_{-1}[S_{AB}] \quad \text{Eq 10}$$

$$\frac{d[S_{AB}]}{dt} = k_1[AB][S_-] - k_{-1}[S_{AB}] \quad \text{Eq 11}$$

Examining the last differential equation as an example, we can substitute concentrations for probability multiplied by total structures:

$$\frac{d[S_{AB}]}{dt} = \frac{d}{dt} (p_{S_{AB}} [S_{tot}]) \quad \text{Eq 12}$$

By the product rule:

$$\frac{d[S_{AB}]}{dt} = \frac{dp_{S_{AB}}}{dt} [S_{tot}] + p_{S_{AB}} \frac{d[S_{tot}]}{dt} \quad \text{Eq 13}$$

We make the simplifying assumption that the total concentration of bound structures does not change throughout the duration of the experiment. Though this may not be strictly true, the loss of bound structures due to denaturation and unbinding is low.

$$\frac{d[S_{tot}]}{dt} = 0 \quad \text{Eq 14}$$

Substituting into Equation 13 dropping the second term and rearranging thus yields an expression for the change in probability of a species :

$$\frac{dp_{S_{AB}}}{dt} = \frac{1}{[S_{tot}]} \frac{d[S_{AB}]}{dt} \quad \text{Eq 15}$$

Substituting with equation 11 we have

$$\frac{dp_{S_{AB}}}{dt} = (k_1[AB][S_-] - k_{-1}[S_{AB}]) \frac{1}{[S_{tot}]} \quad \text{Eq 16}$$

This rate is the same as the net rate of change for a given state in the CTMC model. We can substitute the structure-containing species concentrations for their respective products of probability with total structures.

$$\frac{dp_{S_{AB}}}{dt} = \frac{k_1[AB]p_{S_-}[S_{tot}] - k_{-1}p_{S_{AB}}[S_{tot}]}{[S_{tot}]} \quad \text{Eq 17}$$

Yielding finally an expression of the rate of state change as a function of constituent probabilities that we can utilize as part of the CTMC model.

$$\frac{dp_{s_{AB}}}{dt} = k_1[AB]p_{s_-} - k_{-1}p_{s_{AB}} \quad \text{Eq 18}$$

In order to satisfy the above equation, the units of each term on the right hand side must agree with that of the left T^{-1} (i.e. sec^{-1}). The concentration of free antibodies, which has units of NV^{-1} (i.e. molarity M).

k_1 thus has units of $\text{M}^{-1}\text{sec}^{-1}$ and k_{-1} has units sec^{-1} .

At equilibrium, the following expression holds:

$$[AB][S_-]k_1 = [S_{AB}]k_{-1} \quad \text{Eq 19}$$

We can substitute the concentrations of structure-containing species with the product of their probabilities and the total number of structures.

$$[AB]p_{s_-}[S_{tot}]k_1 = p_{s_{AB}}[S_{tot}]k_{-1} \quad \text{Eq 20}$$

$[S_{tot}]$ can be dropped from the equation yielding:

$$[AB]p_{s_-}k_1 = p_{s_{AB}}k_{-1} \quad \text{Eq 21}$$

The following dimensionless expression can thus be formed by rearranging the above equation, and we can see that the units of k_1 and k_{-1} satisfy the dimensionlessness of the left hand side:

$$\frac{p_{s_-}}{p_{s_{AB}}} = \frac{k_{-1}}{[AB]k_1} \quad \text{Eq 22}$$

We can assemble analogous expressions governing the interconversion between monovalent and bivalent states which do not depend on the concentration of free antibodies. An instance is represented with the chemical equation



Where $[S_{AB-b}]$ represents the concentration of bivalently bound structure-antibody complexes.

For which we also have

$$[S_{tot}] = [S_{AB}] + [S_{AB-b}] \quad \text{Eq 24}$$

We finally arrive at a probabilistic transition rate expression:

$$\frac{dp_{s_{AB}}}{dt} = k_2p_{s_{AB}} - k_{-2}p_{s_{AB-b}} \quad \text{Eq 25}$$

k_2 and k_{-2} both have units of sec^{-1} .

Conversion of SPR response signal to model-compatible quantity.

Earlier we defined occupancy as the number of antibodies associated with a structure as a function of state. A structure with four antigen sites and 1 bivalently bound antibody or a structure with one

antigen site and a single monovalently bound antibody both have an occupancy of 1. A structure with two antibodies bound, regardless of configuration has an occupancy of 2 etc. The average occupancy of all structures is the expected value function of occupancy or the sum of the occupancy of all states weighted by their probability distribution.

$$O = \sum_{i=1}^N p_i o_i \quad \text{Eq 26}$$

where N is the total number of structure states, p_i is the probability of a particular state, and o_i is the occupancy or number of antibodies associated with that state.

We can express SPR signal as a product of a conversion constant and the amount of bound antibody species represented by the average occupancy O multiplied by the total concentration of structures.

$$R = R_{conv}[S_{tot}]O = R_{conv}[S_{tot}] \sum_{i=1}^N p_i o_i \quad \text{Eq 27}$$

Next we design a single calibration experiment for obtaining a conversion factor from SPR response signal to occupancy. Since in a single antigen configuration, the only contributing state to occupancy is S_{AB} , i.e. $O = p_{S_{AB}} o_{S_{AB}}$ and $o_{S_{AB}} = 1$ the following holds true:

$$R = R_{conv}[S_{AB}] \quad \text{Eq 28}$$

and also

$$R = R_{conv} p_{S_{AB}} [S_{tot}] \quad \text{Eq 29}$$

For our calibration, with the single antigen configuration it is possible to perform a saturation step where a high concentration of free antibodies and extended saturation time are used to ensure maximum average occupancy – i.e. where every structure contains a monovalently bound antibody. We can represent such a situation with the following expression:

$$[S_{AB}]_{\max_cal} = [S_{tot}]_{cal} \quad \text{Eq 30}$$

where $[S_{tot}]_0$ is the total number of structures in this calibration experiment.

A probability expression analogous to Equation 4, can be written in the saturation condition:

$$p_{S_{AB}\max_cal} = \frac{[S_{AB}]_{\max_cal}}{[S_{tot}]_{cal}} = 1 \quad \text{Eq 31}$$

And conversely for probabilities to sum to 1, the following must be true:

$$p_{S_{-}\max_cal} = 0 \quad \text{Eq 32}$$

where $p_{S_{-}\max_1ab}$ is the probability of an empty structure at saturation.

The measurable SPR response signal during saturation of the single antigen configuration is

$$R_{\max_cal} = R_{conv}[S_{AB}]_{\max_cal} \quad \text{Eq 33}$$

And since we can rewrite concentrations in terms of their state probability and the total number of structures, we have

$$[S_{AB}]_{\max_cal} = p_{S_{AB}\max_cal}[S_{tot}]_{cal} = [S_{tot}]_{cal} \quad \text{Eq 34}$$

Substituting, we can then solve for the conversion factor:

$$R_{conv} = \frac{R_{\max_cal}}{[S_{AB}]_{\max_cal}} = \frac{R_{\max_cal}}{[S_{tot}]_{cal}} \quad \text{Eq 35}$$

Applying the conversion factor to any measured SPR signal we can thus obtain the average occupancy.

$$O = \frac{R}{R_{conv}[S_{tot}]} = \frac{R}{R_{\max-1ab}} \frac{[S_{tot}]_{cal}}{[S_{tot}]} \quad \text{Eq 36}$$

Where $[S_{tot}]$ varies from experiment to experiment and must thus be obtained independently to normalize each run relative to the number of structures.

Breadth-First Enumeration of States and Transitions

Modeling the kinetics of multivalent Ab binding begins with an enumeration of the possible states where each state consists of the pattern of antigens and the specific arrangement of Ab bound to the pattern.

We approximate a single nanostructure and containing g antigen sites as a set of points in 2 dimensional space.

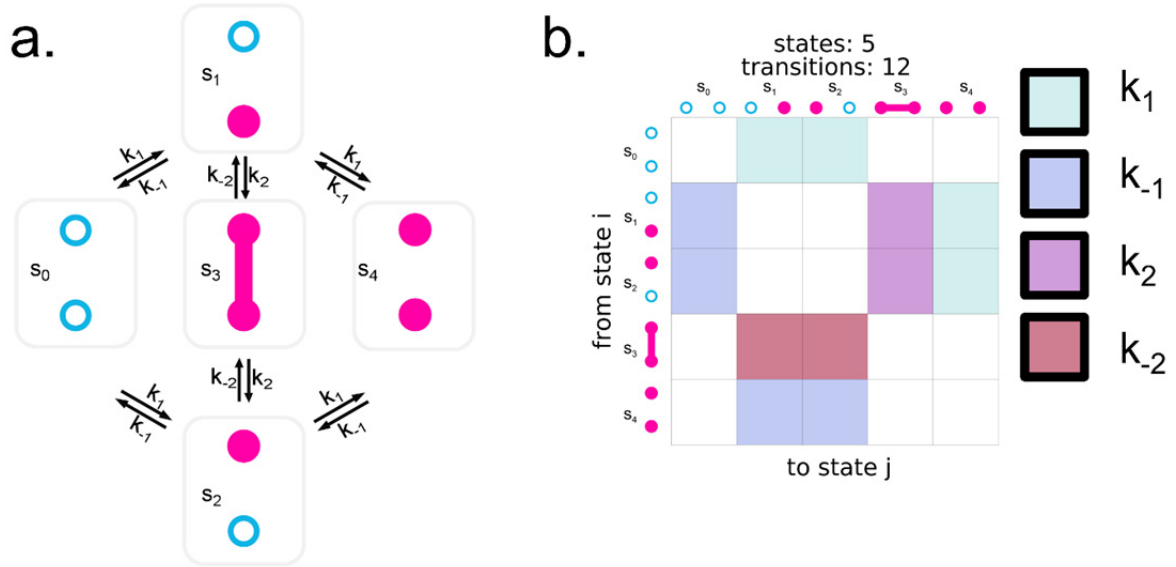
$$X = \{x_0, x_1, \dots, x_{g-1}\} \quad \text{Eq. 37}$$

We are interested in knowing which pairs of antigen sites are eligible for bivalent binding by nature of their proximity to each other.

The symmetric distance matrix \mathbf{D} contains the Euclidian distances D_{ij} for each pair of points in \mathbf{x} . We can treat the antigen sites as an undirected graph with adjacency matrix \mathbf{A} of neighboring sites by applying a critical distance criterion, i.e. the distance d_c beyond which no significant bivalent binding can occur because the sites are too far to be bridged by any antibody. An element A_{ij} has the value.

$$A_{ij} = \begin{cases} 1 & \text{if } D_{ij} < d_c \\ 0 & \text{if } D_{ij} \geq d_c \end{cases} \quad \text{Eq. 38}$$

Our strategy to simulate the dynamics of antibody binding on complex antigen configurations is to decompose the process of binding into elementary steps that comprise a set $k = \{k_1, k_{-1}, k_2, k_{-2}\}$ of elementary rate parameters corresponding respectively to monovalent binding, monovalent unbinding, monovalent to bivalent conversion, and bivalent to monovalent conversion. This enables us to simulate the process as a continuous time Markov chain (CTMC) with a limited number of parameters that are generalizable to many different multivalent antigen configurations. To satisfy the Markov property or memorylessness of a given state, we must have access to the complete set of transitions from every possible state (i.e. thermodynamic microstate) to any existing adjacent state. A given antigen configuration has a unique set of states $S = \{s_0, s_1, \dots, s_{n-1}\}$ each of which corresponds to a unique permutation of empty sites and Abs bound either monovalently or bivalently to the antigen configuration (e.g. **Supplementary Fig. 19**). For the purposes of state recall and monitoring of redundancy, each state is assigned a unique ID upon discovery from which its exact arrangement can be reproduced.



Supplementary Fig. 19. An example of the network of states and transitions for a simple 2-site configuration. A. Gray boxes indicate unique states of the system that are connected by fundamental processes governed by transition rates. Cyan-colored circles indicate empty sites, whereas pink indicate sites with Abs bound. Edges connecting two sites indicate a bivalent connection to two antigens whereas single pink dots indicate monovalently bound antibodies. B. The corresponding transition matrix equivalent to the scheme in A where the vertical axis indicates the origin of a transition and the horizontal axis indicates the destination. The colors of each cell indicate the rate governing a particular transition.

The states are related to each other by transitions that we can represent as a sparse, directed graph with an $n \times n$ transition matrix \mathbf{T} which has transition rate elements.

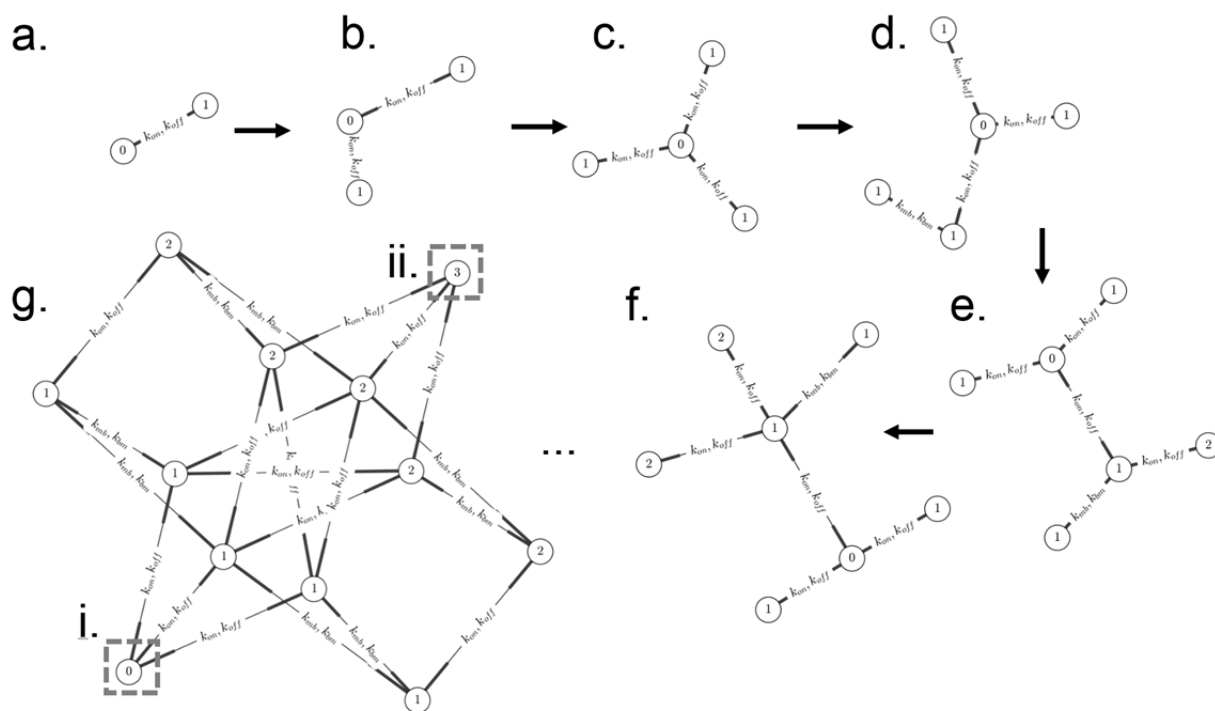
$$R_{i,j} = R(s, s') = \begin{cases} k(i, j) & \text{if } A_{ij} = 1 \\ 0 & \text{if } A_{ij} = 0 \end{cases} \quad \text{Eq 39}$$

where $k(i, j) \in \{k_1, k_{-1}, k_2, k_{-2}\}$ and $R_{s,s'}$ is the rate of leaving state s to go to state s' .

For example, the 2-site configuration in Suppl. Fig 17A has transition rate matrix.

$$\mathbf{R} = \begin{pmatrix} 0 & k_1 & k_1 & 0 & 0 \\ k_{-1} & 0 & 0 & k_2 & k_1 \\ k_{-1} & 0 & 0 & k_2 & k_1 \\ 0 & k_{-2} & k_{-2} & 0 & 0 \\ 0 & k_{-1} & k_1 & 0 & 0 \end{pmatrix}. \quad \text{Eq 40}$$

Legal state transitions occur through site operations i.e. changes in the number of Abs and the valence of their binding (mono or bi). Though many unique states and transitions can occur for a given configuration, they occur through the limited number of elementary processes. For example in the 2-site configuration represented in **Supplementary Fig. 19**, transitions $R_{0,1}$, $R_{0,2}$, $R_{2,4}$, and $R_{1,4}$ have the same rate k_1 .

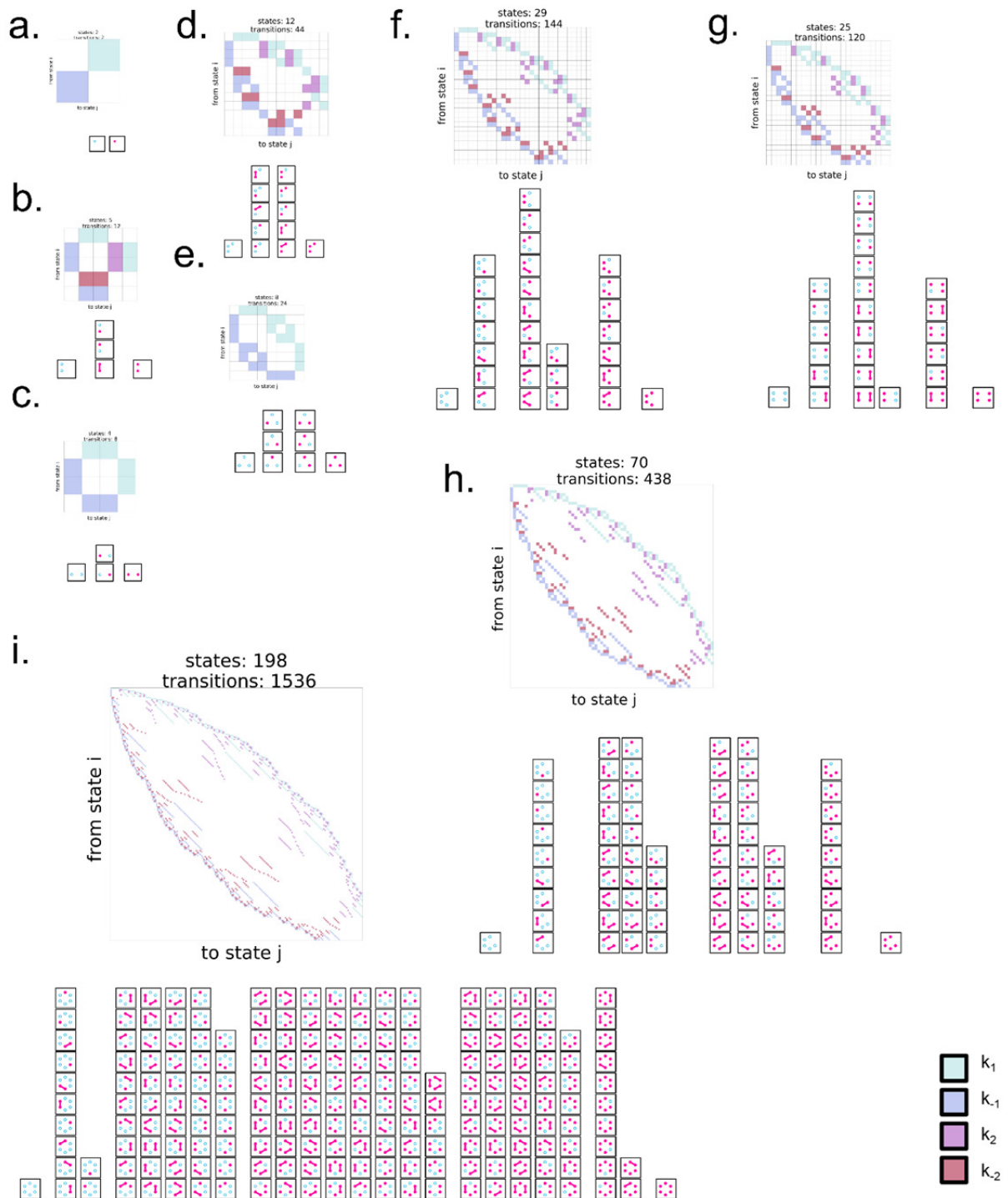


Supplementary Fig. 20. Visualization of the breadth first search as the network of states and transitions is constructed through systematic neighbor-to-neighbor visitation of individual states connected by transitions via the four fundamental processes. Edge labels indicate forward and reverse rates governing a particular transition whereas nodes indicate specific states and node labels indicate the number of bound antibodies in that state. (A-F) Consecutive steps of the BFS. (G) The final network at the conclusion of the BFS with dashed-line boxes (i) and (ii) indicating peripheral states with 0 and 3 (the maximum for this system) bound antibodies respectively.

We implemented a breadth-first search (BFS) algorithm for enumerating the states and transitions for an arbitrary set of antigen points. The algorithm both discovers the existence of edges between neighboring states as well as the type of process taking place and assigns the corresponding rate (**Supplementary Fig. 20**). BFS works by creating a queue of neighboring nodes upon visiting a node and the subsequently visiting each node in that queue, working recursively until all nodes with a path to the start have been visited. In this case, each node corresponds to a state of the antigen-antibody system, and a neighbor state is discovered through physical changes to the system, for example by attempting to add an Ab monovalently to an empty site or converting an existing bivalent Ab_y back to a monovalent state consisting of two sites, one empty the other a monovalent complex. Exploring this state space is trivial and can be accomplished by a human for simple cases such as the two site example in **Supplementary Fig. 20**. However for complex systems, the number of possible states and connections between those states becomes large demanding more automated enumeration. The resulting transition matrix is then be used to simulate the evolution of the system starting from an initial condition.

Supplementary Fig. 21. shows visual representations of the transition matrix and the complete set of visual representations of each state organized according to number of Abs (occupancy) corresponding to each particular pattern among the various partial and complete hexameric geometries examined in this study. We can see here how the more interconnected states have a

much larger set of possible states. Two 2-site configurations have different separation distances, one within the 21 nm range of possible bivalent connectivity (bivalent close **Supplementary Fig. 2b**) and one where the two antigens are separated far enough from each other to prohibit bivalent connectivity (**Supplementary Fig. 21c**) resulting the former having a greater number of possible states and transitions 5 and 12 respectively versus 4 and 8 respectively. Similar comparisons can be made with 3 site configurations such as the obtuse connected triangle (**Supplementary Fig. 21d**) with 12 states and 44 transitions versus the separated equilateral triangular configuration (**Supplementary Fig. 21e**) with only 8 states and 24 transitions. A final comparison can be made with 4-site configurations: connected trapezoidal (**Supplementary Fig. 21f**) with 29 states and 144 transitions and the partially separated rectangle (**Supplementary Fig. 21g**) with 25 states and 120 transitions.

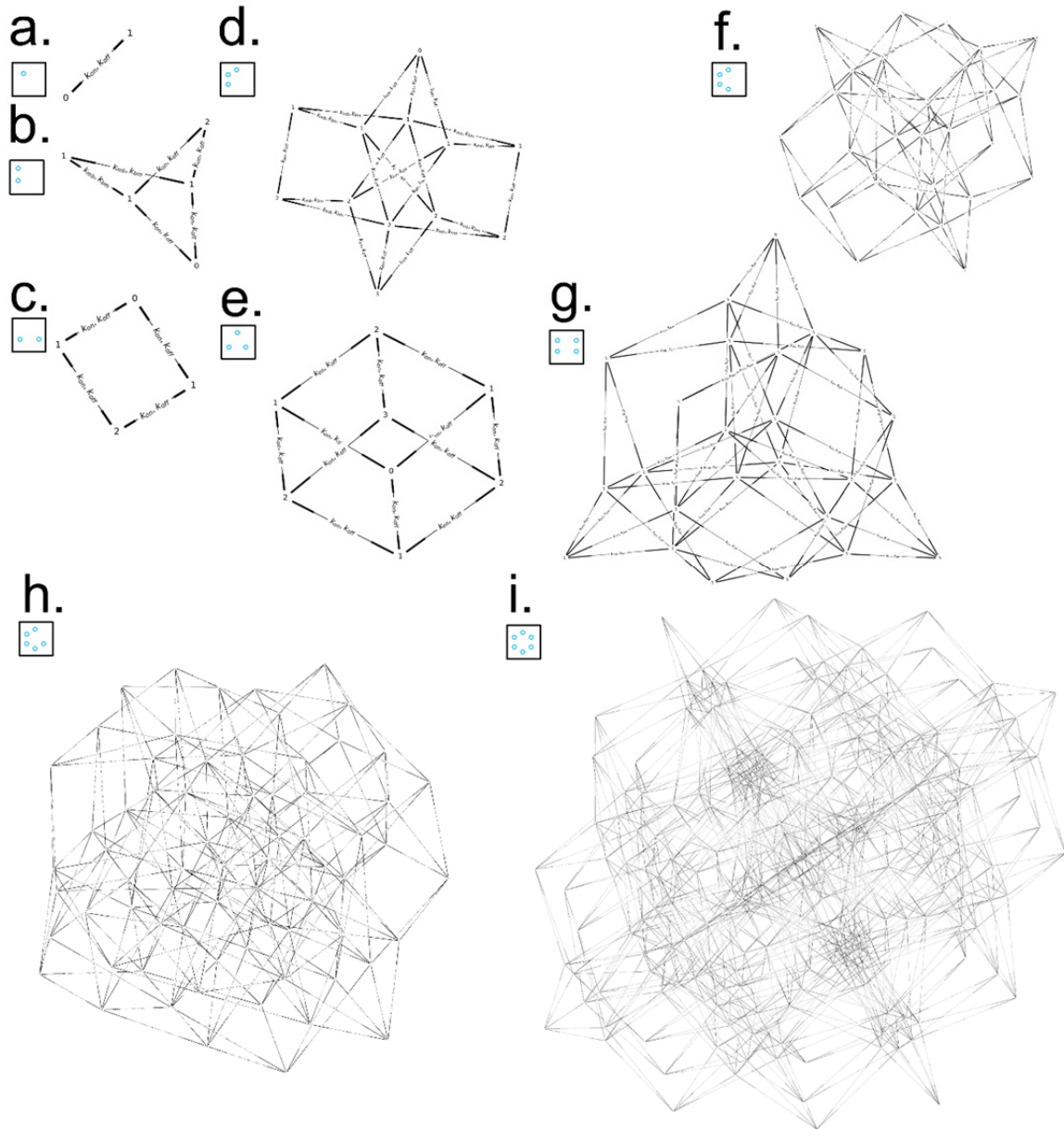


Supplementary Fig. 21. Transition matrices and visualizations of states identified through the breadth first search algorithm for each of the configurations for the antigen hexamer. Colors indicate which of the 4 fundamental processes governs a particular transition from state i (vertical axis) to state j (horizontal axis). a. monovalent configuration b. bivalent (close) configuration c. bivalent (far) configuration d. obtuse triangle configuration e. equilateral triangle (far) configuration f. trapezoid configuration g. rectangle (far) configuration h. pentagon configuration i. hexagon configuration.

Due to the reversibility of each of the fundamental processes, topologically none of the networks contain sinks or sources. Furthermore, the redundant use of fundamental processes leads to

symmetries, for example monovalent binding to site 1 in a 2-site configuration is mirrored by monovalent binding to site 2 in the 2-site configuration, creating a symmetrical branch to two degenerate states. Topology alone gives us some indications as to the relationships between states as well as an appreciation for the necessity of traversing certain states in order to reach others. Visualizations of the networks for the various partial and complete hexameric configurations are shown in **Supplementary Fig. 22**.

Even in complex systems such as the complete hexameric configuration exhibited striking visual symmetry due to the degenerate nature of many different states, for example there are 24 ways to arrange 1 monovalently bound Ab and 1 bivalently bound antibody in the hexameric configuration that are all functionally and topologically isomorphic if there is equal spacing between vertices of the hexagon.



Supplementary Fig. 22. Visualizations of the networks of states and transitions for each of the configurations of the antigen hexamer patterns. Edges are bidirectional, indicating both forward and reverse processes eg. k_1 and k_{-1} or k_2 and k_{-2} . a. monovalent configuration b. bivalent (close) configuration c. bivalent (far) configuration d. obtuse triangle configuration e. equilateral triangle (far) configuration f. trapezoid configuration g. rectangle (far) configuration h. pentagon configuration i. hexagon configuration.

Solving the Markov Chain (Deterministically) from the completed state transition matrix

A fully mapped network of states and transitions enables us to reproduce the kinetics of binding by simulating the traversal of a system through the network of states from an initial condition.

The probability vector $\vec{\pi}(t)$ is the time dependent probability distribution of the state space.

$$\vec{\pi}(t) = \begin{pmatrix} \pi_{s_0}^t \\ \pi_{s_1}^t \\ \vdots \\ \pi_{s_{n-1}}^t \end{pmatrix} \quad \text{Eq 41}$$

A uniform distribution $\vec{\pi}(t)$ for an instance of time would imply that at that instant, there is an equal probability of a structure to occupy any of the possible states. Another example is the initial condition at the start of an SPR experiment which is 1 for the state corresponding to an empty structure and 0 for all other states.

$\vec{\pi}(t)$ can be computed via a generalization of the exponential distribution for a matrix of transition rates, the infinitesimal generator matrix which is related to the transition rate matrix by.

$$Q_{i,j} = Q(s,s') = \begin{cases} R(s,s') & s \neq s' \\ -\sum_{s \neq s'} R(s,s') & \text{otherwise} \end{cases} \quad \text{Eq 42}$$

The transient probabilities are then equal to the following matrix exponential.

$$\vec{\pi}(t) = \vec{\pi}(0) \cdot e^{\mathbf{Q} \cdot t} = \vec{\pi}(0) \cdot \sum_{i=0}^{\infty} \frac{(\mathbf{Q} \cdot t)^i}{i!} \quad \text{Eq 43}$$

which is valid and computable, although numerical instability has been reported for this case. The solution is to use the method of uniformized discrete time markov chain, where the continuous chain is approximated by a discrete one by renormalizing the matrix with respect to the fastest outgoing rate.

$$q \geq \max_i \{|Q(i,i)|\} \quad \text{Eq 44}$$

to form a uniformized transition matrix .

$$\mathbf{U} \stackrel{\text{def}}{=} \mathbf{I} + \frac{\mathbf{Q}}{q} \quad \text{Eq 45}$$

that can be used to compute the transient probability distribution of the next time step given a current distribution. The process can be used to begin with an initial condition followed by a series of probability estimates for each subsequent timestep after the initial condition.

The matrix exponential then becomes.

$$\vec{\pi}(t) = \vec{\pi}(0) \cdot e^{\mathbf{Q}t} = \vec{\pi}(0) \cdot e^{q(\mathbf{U}-\mathbf{I})t} = \vec{\pi}(0) \cdot e^{qt\mathbf{U}} \cdot e^{-qt} = \vec{\pi}(0) \cdot e^{-qt} \cdot e^{qt\mathbf{U}} \quad \text{Eq 46}$$

A Taylor series expansion of $e^{qt\mathbf{U}}$ yields the following which is computed iteratively in approximation of the infinite summation.

$$\vec{\pi}(t) = \vec{\pi}(0) \cdot e^{-qt} \cdot \sum_{i=0}^{\infty} \frac{(qt\mathbf{U})^i}{i!} = \sum_{i=0}^{\infty} \frac{e^{-qt}(qt)^i}{i!} \cdot \vec{\pi}(0) \cdot \mathbf{U}^i \quad \text{Eq 47}$$

Thus given a set of transitions, their rates which depend on the binding kinetics of the antibody/antigen complex, and an initial condition, the complete binding and unbinding dynamics of a population of patterned antigens in the presence of antibodies can be simulated.

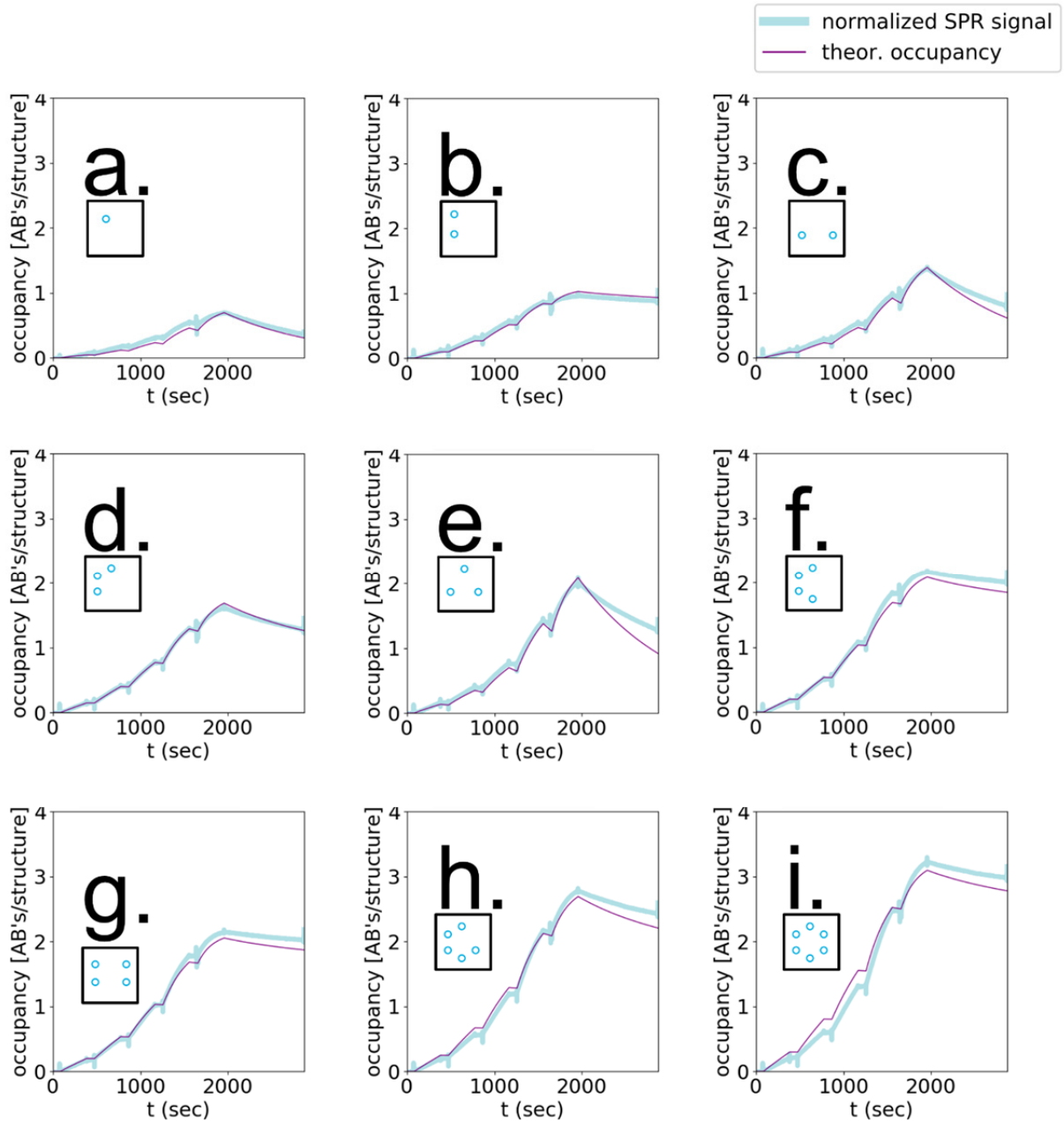
Fitting and experimental calibration

Simulated SPR data with signal proportional to the amount of bound material (total occupancy) can be produced by counting the number of Abs bound (occupancy) and summing this quantity for each state weighted according to the probability of those states. The SPR signal is thus an expected value function of occupancy for a given state probability distribution and a function of time given by.

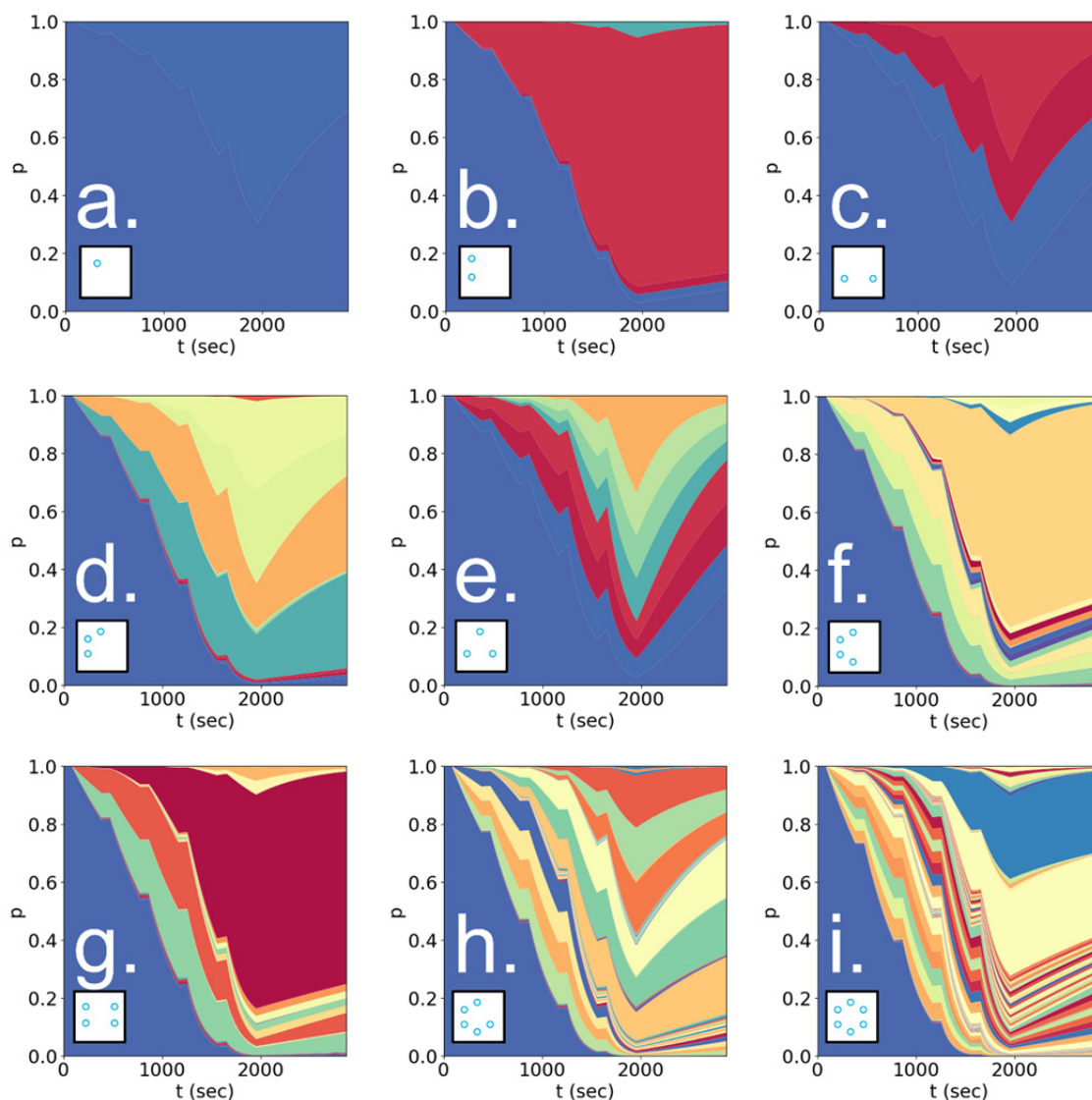
$$O(t) = \sum_{i=1}^N p_i(t) o_i \quad \text{Eq 48}$$

where $O(t)$ is the total occupancy of the system at an instance in time, analogous to the SPR signal, $p_i(t)$ is the probability of a particular state as a function of time, o_i is the integer value of the occupancy of state i which could be for example 2 Abs bound, and the N is the total number of states possible for the pattern.

Since the state probability distribution is known for each time point of a run, the expected distribution of bound material is also known. The values for the four rate parameters can be estimated by minimizing the sum squared error between the simulated and experimental SPR data where the experimental data is normalized to convert SPR signal to an ensemble measurement of the number of Abs bound per structure. This conversion factor is determined empirically with a saturation experiment involving a simple 1-site, monovalent binding scheme at high antibody concentrations to produce the maximum signal achievable corresponding to 1 antibody per structure.

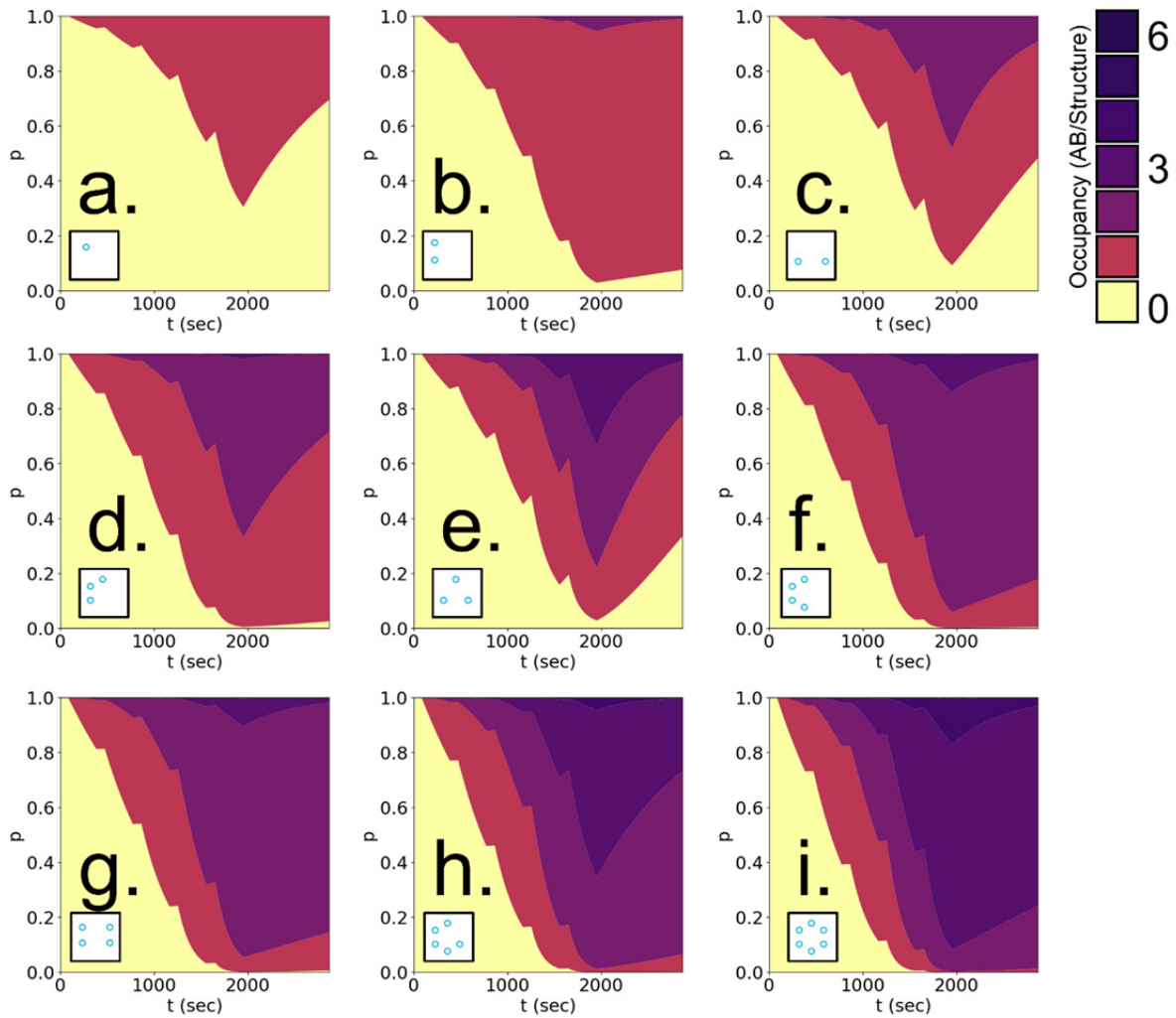


Supplementary Fig. 23. CTMC-model fits to the various partial configurations of antigen hexamer patterns. The cyan line indicates physical data from SPR experiments and magenta indicates the trajectory of the model occupancy curve. A single parameterization was determined for all configurations, rather than best fit for individual configurations, thus the rates are the same across conditions. a. monovalent configuration b. bivalent (close) configuration c. bivalent (far) configuration d. obtuse triangle configuration e. equilateral triangle (far) configuration f. trapezoid configuration g. rectangle (far) configuration h. pentagon configuration i. hexagon configuration. Fitting and simulation was performed once after an iterative optimization process.



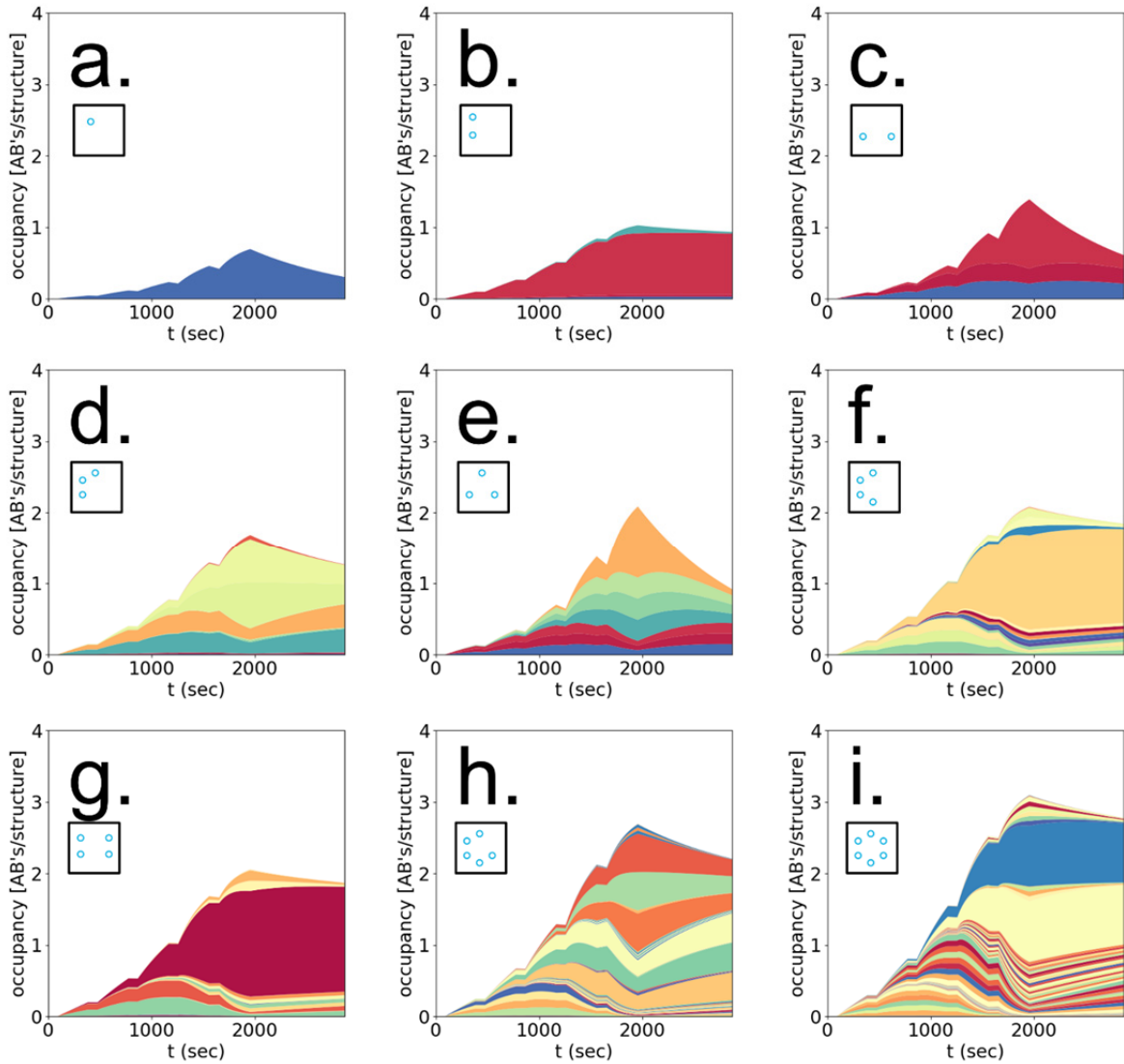
Supplementary Fig. 24. CTMC-model stacked probability trajectories for every state of the various configurations of the antigen hexamer patterns using the fitted rate parameters. Colors are used to distinguish different states. State probabilities satisfy the normalization condition, summing to 1 at every point. a. monovalent configuration b. bivalent (close) configuration c. bivalent (far) configuration d. obtuse triangle configuration e. equilateral triangle (far) configuration f. trapezoid configuration g. rectangle (far) configuration h. pentagon configuration i. hexagon configuration.

The probability distribution can also be represented as a trajectory, with colors used to distinguish different states. This is shown for each of the hexameric configurations in **Supplementary Fig. 24**. Note the steady elimination of the empty state (indicated in blue) which initially constitutes 100% of the probability distribution but which quickly becomes replaced by states of varying occupancy. Note to the acceleration of changes in state distribution due to the increasing concentrations injected at each of the time points of the SPR series.



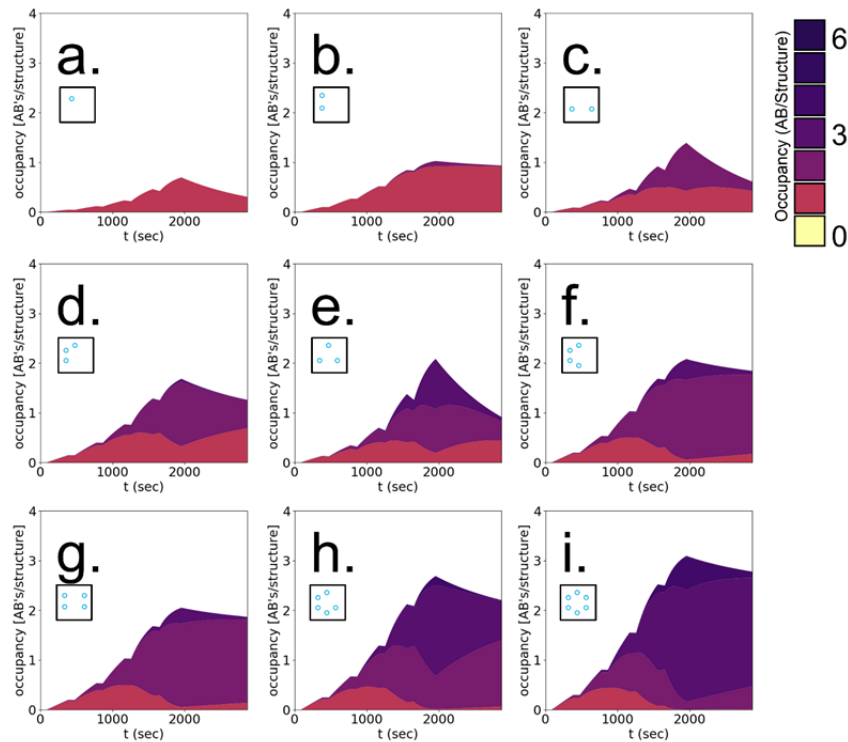
Supplementary Fig. 25. Stacked probability trajectories with states aggregated into macrostates according to their occupancy value. Thus all states with, for example, 2 Abs, regardless of their position or monovalent/bivalent status are considered together and likewise for 0, 1, 3, etc Abs. Color key indicates number of Abs per macrostate with 6 being the largest for the hexameric patterns. a. monovalent configuration b. bivalent (close) configuration c. bivalent (far) configuration d. obtuse triangle configuration e. equilateral triangle (far) configuration f. trapezoid configuration g. rectangle (far) configuration h. pentagon configuration i. hexagon configuration.

It can be instructive to classify the different states according to their occupancy. Multiple states with 1 Ab bound for instance are considered together whether or not they are monovalently or bivalently bound and regardless of their positioning in the pattern. **Supplementary Fig. 25.** shows stacked probability trajectories for these aggregate macrostates.



Supplementary Fig. 26. CTMC-model stacked probabilities weighted by occupancy. The probability of each state weighted according to its occupancy equals its contribution to the SPR signal, and the sum of all weighted probabilities reproduces the SPR curve. Colors distinguish the different states. a. monovalent configuration b. bivalent (close) configuration c. bivalent (far) configuration d. obtuse triangle configuration e. equilateral triangle (far) configuration f. trapezoid configuration g. rectangle (far) configuration h. pentagon configuration i. hexagon configuration.

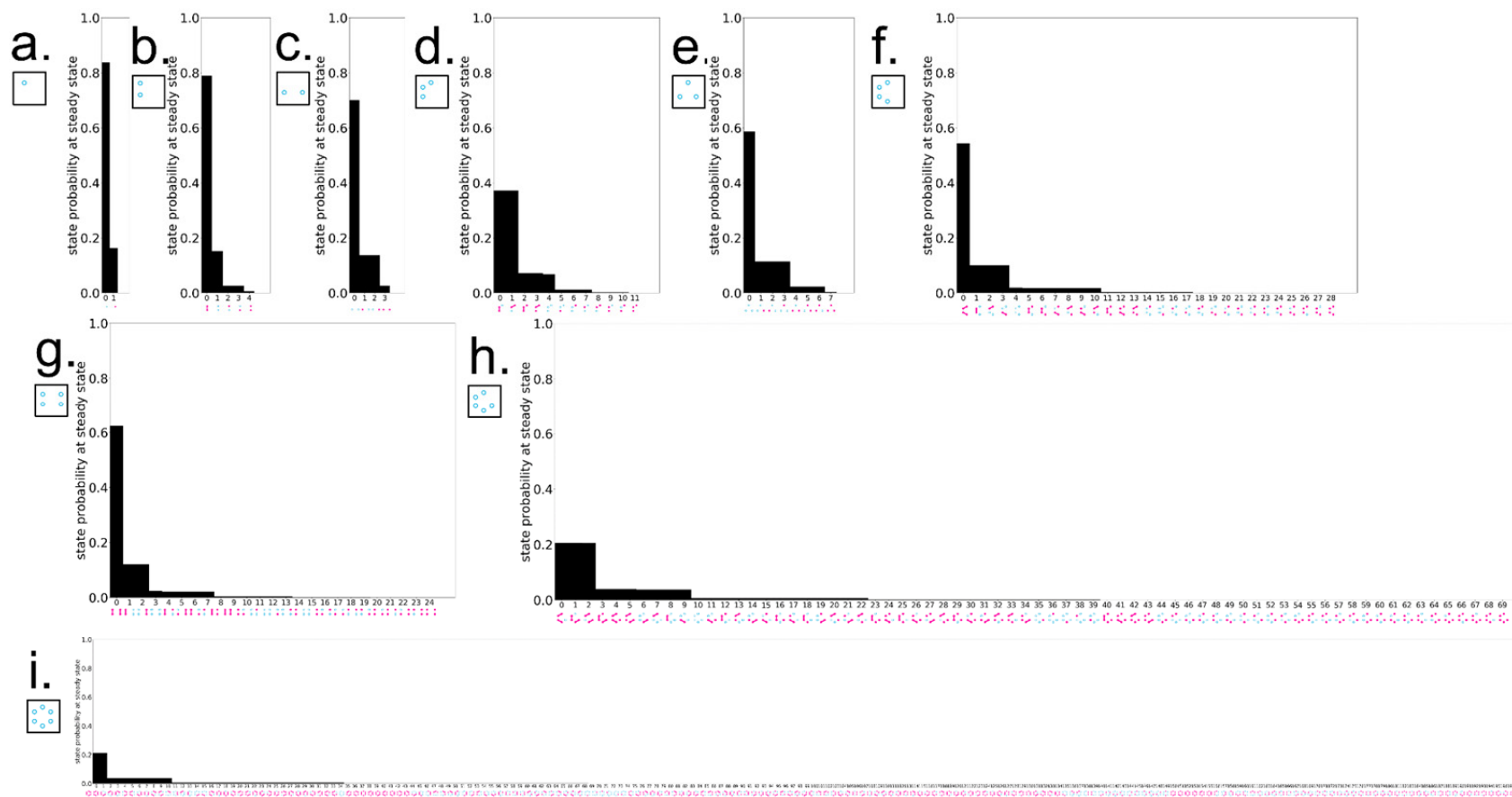
The probability distribution weighted according each state's particular occupancy, described by Eq 48, is visualized as a stacked trajectory in **Supplementary Fig. 26**. Here we see that the stacking of weighted probabilities reproduces the total occupancy at each instance in time, regenerating the SPR curves that were fitted to the experimental data but with added stratification indicating the contributions to this signal from each state. Note in this case, the absence of an empty state contribution, as the occupancy is zero for this particular state, making its contribution to the total occupancy zero as well.



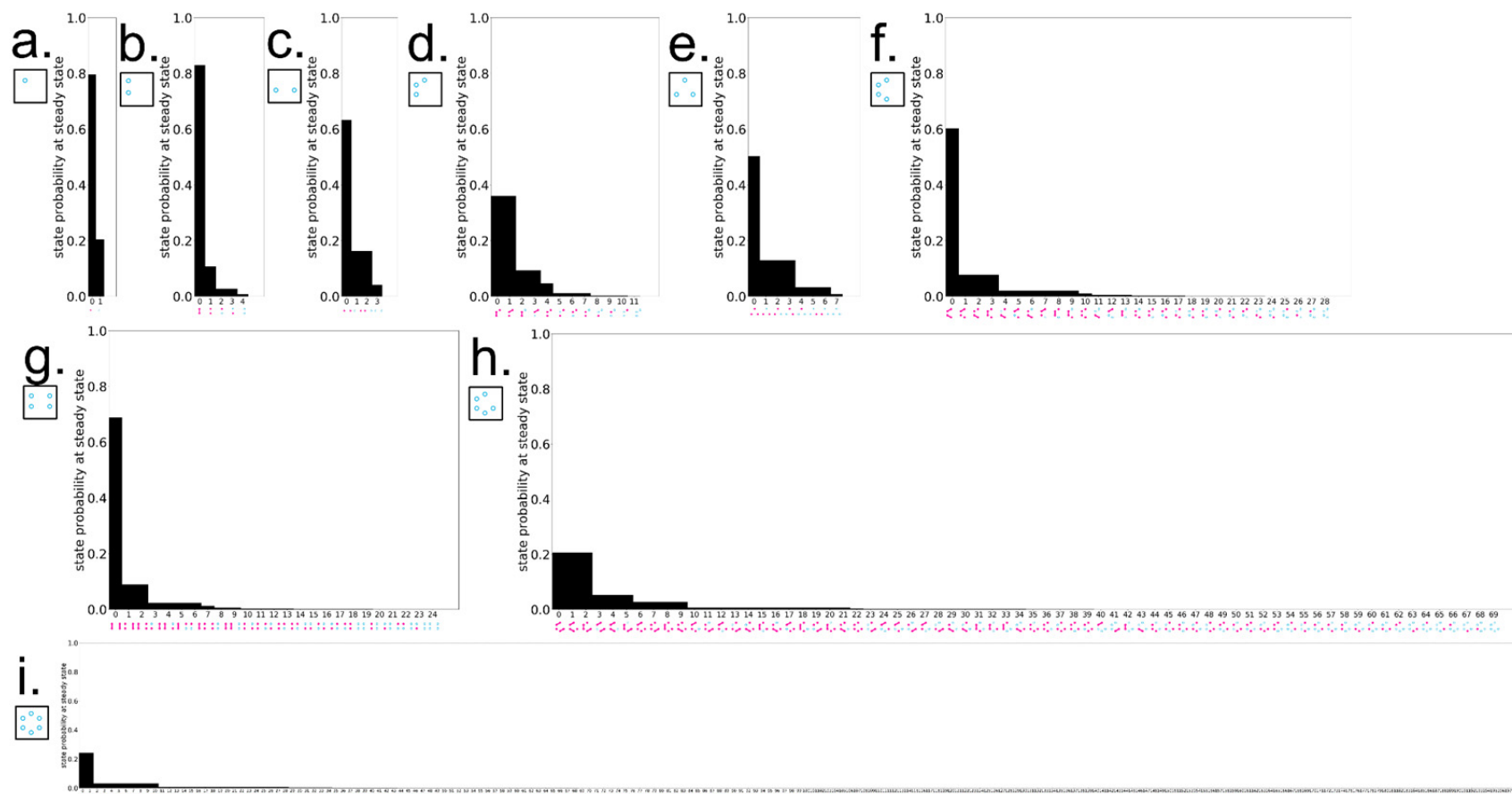
Supplementary Fig. 27. Stacked probabilities weighted by occupancy and aggregated according to occupancy of each state to form macrostates. Colors distinguish the different macrostates with 6 antibodies being the largest for the hexameric configuration. a. monovalent configuration b. bivalent (close) configuration c. bivalent (far) configuration d. obtuse triangle configuration e. equilateral triangle (far) configuration f. trapezoid configuration g. rectangle (far) configuration h. pentagon configuration i. hexagon configuration.

An occupancy-aggregated macrostate version of the weighted stacked probability trajectories is shown in **Supplementary Fig. 27**. Here we can appreciate the dominance or lack thereof of certain macrostates which are, in cases where bivalent binding is strongly favored, much less than the maximum possible occupancy – for example in the case of the complete hexagon which is dominated by an occupancy of 3 despite having a maximum possible occupancy of 6 Abs.

At the endpoint of these runs, the distribution of states is not uniform, and in fact certain states dominate due to a combination of network topology and transition rate-dependent conduction of flow through the network of states. Of particular interest is the steady state behavior of these distributions, and we estimate this through simulation out to 10000 seconds for a given solution phase Ab concentration and empty structure initial condition. The histogram of state probabilities organized according to rank of each state for a low 0.025 nM concentration is shown in **Supplementary Fig. 28** and for a high 0.5 nM concentration in **Supplementary Fig. 27**. The top ranking states for the different conditions is instructive, and we can see that, for example in the case of the pentameric configuration (**Supplementary Fig. 28h** and **29h**), different top-ranked states emerge depending on the concentration. In contrast, the complete hexameric configuration (**Supplementary Fig. 28i** and **S29i**) yields the same top 2 ranked states in both cases.



Supplementary Fig. 28. Histograms of the state probability distributions according to rank. Horizontal axis indicates state rank with visual depiction of the state. Distributions represent estimates of the steady state distribution after 10000 seconds equilibration in the presence of 0.025 nM concentration of solution phase Abs. a. monovalent configuration b. bivalent (close) configuration c. bivalent (far) configuration d. obtuse triangle configuration e. equilateral triangle (far) configuration f. trapezoid configuration g. rectangle (far) configuration h. pentagon configuration i. hexagon configuration.



Supplementary Fig. 29. Histograms of the state probability distributions according to rank. Horizontal axis indicates state rank with visual depiction of the state. Distributions represent estimates of the steady state distribution after 10000 seconds equilibration in the presence of 0.5 nM concentration of solution phase Abs. a. monovalent configuration b. bivalent (close) configuration c. bivalent (far) configuration d. obtuse triangle configuration e. equilateral triangle (far) configuration f. trapezoid configuration g. rectangle (far) configuration h. pentagon configuration i. hexagon configuration.

Supplementary Note 3

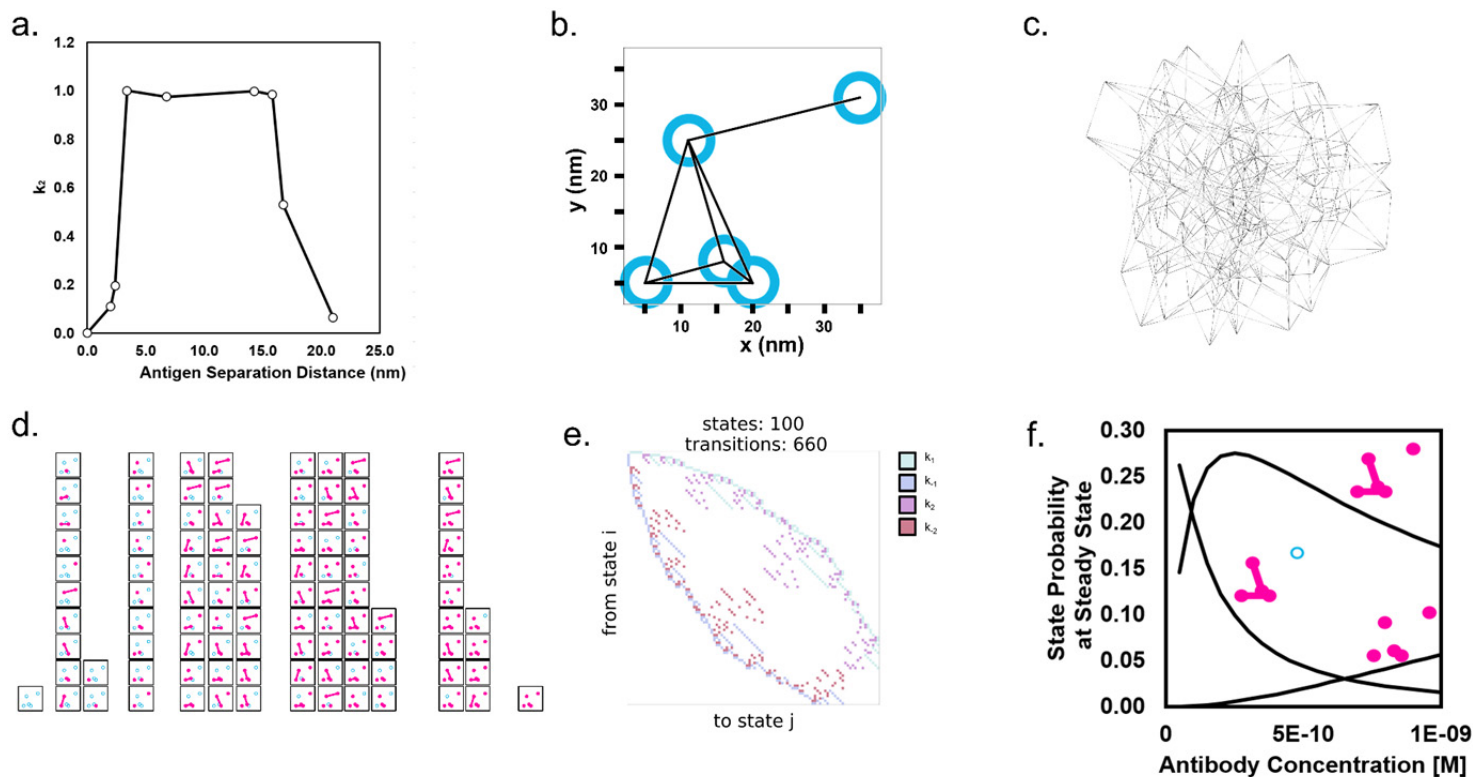
Incorporation of rate dependence on distance and dealing with arbitrary separation distances.

We performed fitting experiments to bivalent configurations (2 antigens) for multiple separation distances of the digoxigenin antigens with IgG Ab as well as NIP antigens and IgG Abs. We make the simplifying assumption that monovalent on-rate (k_1), monovalent off-rate (k_{-1}), and the bivalent to monovalent conversion (k_{-2}) are unaffected by separation distance and fixed these values to their optima that had been determined in the hexameric fitting experiments. Two parameter fits were performed with variable monovalent to bivalent conversion rate (k_2) as well as variable normalization constant (the conversion factor relating SPR signal to a number of Abs bound per structure) and observed the distance dependence curve shown in **Supplementary Fig. 30 a**.

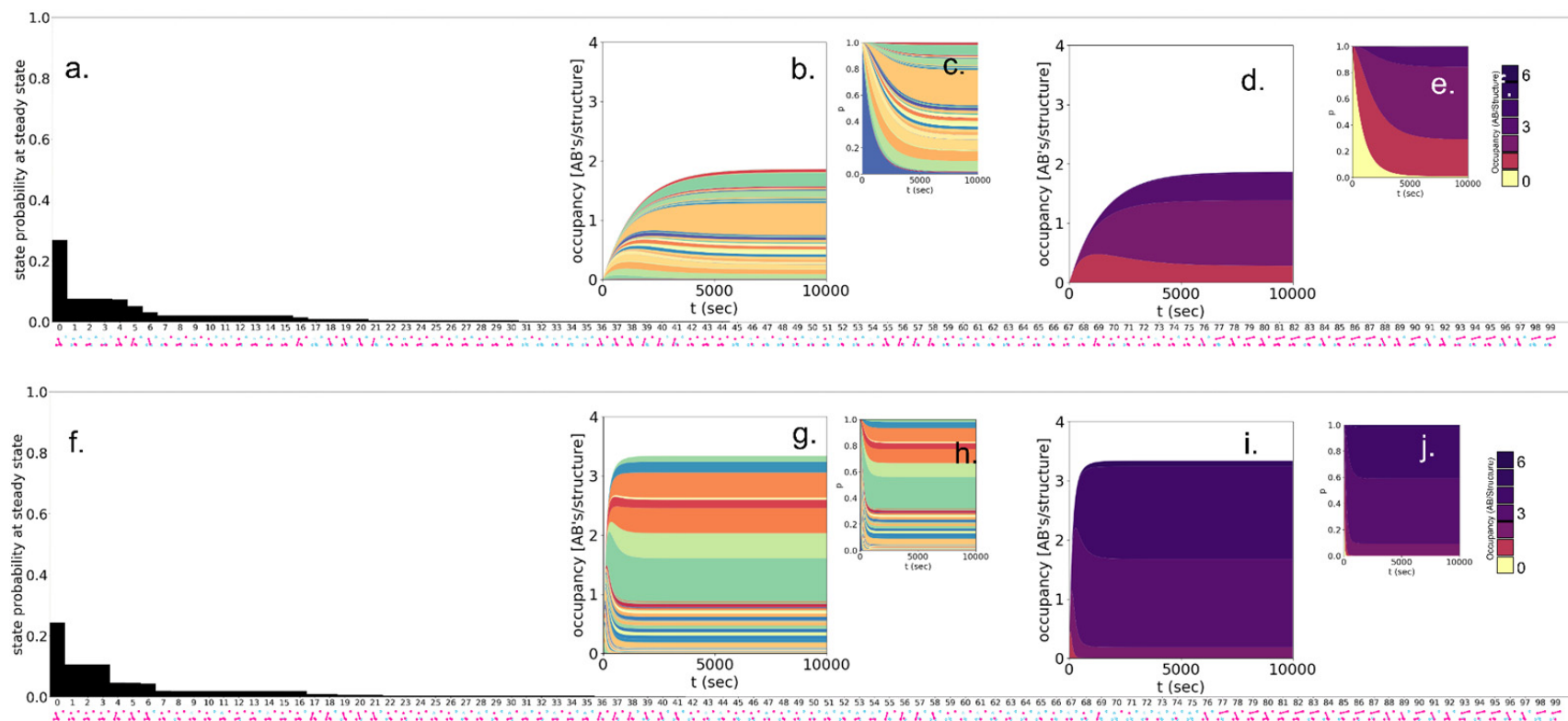
The distance dependence curve was used as a piecewise function to determine what percent of the maximum k_2 value should be used to govern the rates connecting states related by a monovalent to bivalent conversion event. Depending on the distance between the separated antigens, the rate should be less than or equal to the maximum value which corresponds to the ideal separation distance. This region was experimentally confirmed to occupy a band between 6.8 and 17 nm, and distances outside of this range exhibit quickly attenuating conversion rates. We could thus model the dynamics of a complex set of arbitrarily positioned antigens shown in **Supplementary Fig. 30 b-e**.

This procedure was unnecessary for the hexameric configurations as all bivalent distances were either an ideal 15 nm separation or outside the range that antibodies were experimentally observed to reach. The analysis of the arbitrary configuration interestingly shows how concentration dependent pattern preferences can emerge. We performed steady state analysis at multiple concentrations and tracked the probability of specific states that changed relative to one another as different concentrations were used (**Supplementary Fig. 30 f**).

A summary of the simulation data obtained for the arbitrary configuration at near-steady state (10000 seconds) for low 0.025 nM and high 0.5 nM concentrations is shown in **Supplementary Fig. 31**. In the ranking of states by probability (**Supplementary Fig. 26 a and f**) we can see how states with the same occupancy can be favored or disfavored according to their separation distances, with states that include a bivalent connection between the upper-most pair of antigens being highly disfavored.



Supplementary Fig. 30. Summary of experimental calibration and implementation of distance dependence in the multivalent Ab binding model. A. Experimental best-fit results of k_2 to produce a distance dependence function where each point is indicative of the k_2 value for a particular SPR experiment, the horizontal axis indicating the separation distance of the antigens in that experiment, and the vertical axis indicating the fraction of the maximum k_2 value. B. The configuration for a set of arbitrarily chosen antigens whose edges span the complete range of too-close, just-right, and too-far regimes of the distance dependence curve from A. C. Visualization of the network of states and transitions for the arbitrary configuration. D. The set of all states for the arbitrary configuration with a threshold distance of less than 21 nm required to form a bivalent connection. E. Transition matrix for the arbitrary configuration. F. Phase diagram of the estimated steady state probability of selected states (depicted visually) as a function of the solution phase antibody concentration. Note that this model does not take into account possible steric hindrance effects between monovalently bound Abs that could occur at close separation distances.



Supplementary Fig. 31. Summary of steady state analysis results for the arbitrary configuration at low 0.025 nM solution phase antibody concentration (top) and high 0.5 nM concentration (bottom). A, F Histograms of ranked state probabilities. B, G Stacked weighted probability trajectories. C, H Stacked unweighted probability trajectories. D, I Stacked weighted occupancy macro state trajectories. E, J Stacked unweighted occupancy macro state trajectories.

Supplementary Note 4: Rationalization of our choice of antigen–antibody pairs.

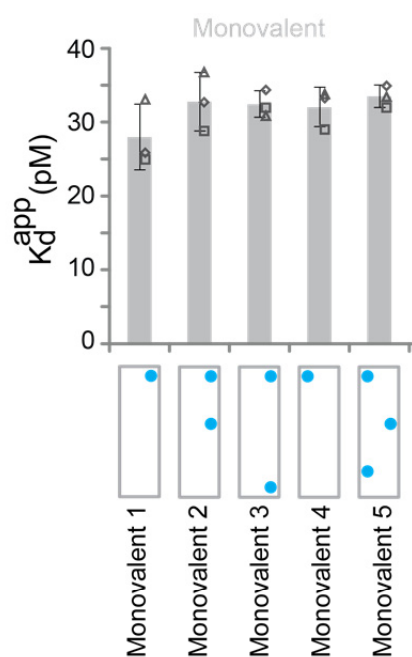
Human anti-NIP IgGs: Most medically relevant antibodies to study, but they cannot be reliably applied to study complex binding patterns in Fig. 4 since their monovalent binding affinity to NIP is weak and reaches steady state rapidly.⁶² This means that kinetic information cannot be extracted from binding curves.

Rabbit anti-DIG IgGs: Less medically relevant but its monovalent binding to DIG still exhibits exponential binding behavior, therefore, kinetic information can be extracted from the binding curves⁶², allowing us to calibrate our markov chain model.

Supplementary Table 2. Binding kinetic data for digoxigenin nanopatterns. (n = 3, the average value of KD and their corresponding standard deviation is presented)

Sample	KD (M)	STDEV for KD (M)
Monovalent 1	2.79E-11	4.45E-12
Bivalent 1	2.08E-12	2.30E-13
Bivalent 2	2.04E-12	2.71E-13
Monovalent 2	3.27E-11	3.96E-12
Monovalent 3	3.24E-11	1.79E-12
Monovalent 4	3.20E-11	2.66E-12
Triangle 1	1.41E-11	1.28E-12
Triangle 2	1.59E-11	2.32E-12
Triangle 3	1.23E-11	1.39E-12
Triangle 4	1.31E-11	3.10E-13
Bivalent 3	3.73E-12	2.56E-13
Monovalent 5	3.35E-11	1.51E-12

Supplementary Table 3. One-way ANOVA analysis of monovalent binding samples. (n = 3)



Fitted binding curve data	Kd Repeat 1	Kd Repeat 2	Kd Repeat 3			
Monovalent 1	2.585E-11	2.493E-11	3.305E-11			
Monovalent 2	3.27E-11	2.88E-11	3.671E-11			
Monovalent 3	2.095E-11	2.318E-11	2.941E-11			
Monovalent 4	3.324E-11	2.895E-11	3.382E-11			
Monovalent 5	3.068E-11	2.794E-11	3.10333E-11			
ANOVA						
<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	1.3244E-22	4	3.31091E-23	2.563317	0.103623	3.47805
Within Groups	1.2916E-22	10	1.29165E-23			
Total	2.616E-22	14				

Supplementary Table 4. Binding kinetic data for NIP nanopatterns using human IgG and IgM and IgG3 hinge engineered variants. The nanostructures are described in Supplementary Fig. 5. The monovalent binding samples (Mono; 28 nm; 45 nm) are weak binding interactions. This means that antibody binding to the antigen coated SPR surface reaches instant equilibrium, and kinetic rate constants cannot be accurately extracted from the binding curves, therefore only the binding constant (KD) is presented. (n = 3 for all bivalent binding samples besides: IgG1 7 nm, IgG2 3 and 7 nm, IgG4 7, 14 and 16, IgM Mono 14, 16 and 17 nm which have 4 repeats; n = 1 for monovalent binding samples (Mono, 28 nm and 45 nm))

Sample_IgG1									Sample_IgG3_HMS								
Sample_IgG1	ka (1/Ms)	Stdev	kd (1/s)	Stdev	KD (M)	Stdev	Chi2	Stdev	Sample_IgG3_HMS	ka (1/Ms)	Stdev	kd (1/s)	Stdev	KD (M)	Stdev	Chi2	Stdev
Mono					3.41E-08		0.0528		Mono					3.94E-08		0.0127	
3 nm	1.53E+07	1130796.18	4.33E-04	9.59523 E-05	2.81E-11	4.42579 E-12	6.31E-02	0.011868	3 nm	1.17E+07	1192741.939	2.20E-04	1.07156 E-05	1.90E-11	2.38395 E-12	7.04E-02	0.00738
7 nm	1.63E+07	920326.0292	1.25E-04	2.28861 E-05	7.69E-12	1.13065 E-12	5.24E-02	0.016642	7 nm	1.26E+07	1118808.89	2.87E-04	1.11447 E-05	2.28E-11	1.63812 E-12	5.68E-02	0.007661
14 nm	2.99E+07	7997164.081	1.07E-04	3.01878 E-05	3.58E-12	1.65756 E-13	3.38E-02	0.008864	14 nm	2.54E+07	5807515.246	1.42E-04	7.7288 E-05	6.17E-12	3.9552 E-12	5.48E-02	0.015946
16 nm	4.74E+07	3376571.239	5.84E-05	1.61931 E-05	1.22E-12	2.52865 E-13	3.96E-02	0.0137	16 nm	3.82E+07	6321157.594	6.65E-05	3.18372 E-05	1.86E-12	1.0623 E-12	4.01E-02	0.008603
17 nm	7.42E+06	322678.3744	3.26E-04	5.31502 E-05	4.42E-11	8.45653 E-12	1.06E-01	0.010013	17 nm	5.98E+06	49369.35622	4.17E-04	6.19866 E-06	6.97E-11	1.56758 E-12	9.85E-02	0.010443
28 nm							9.23E-09	2.47E-01	28 nm	3.07E+05				1.19E-09		3.57E-01	
45 nm							4.05E-08	5.82E-02	45 nm					3.68E-08		3.30E-02	
Sample_IgG2									Sample_IgG3_m15								
Sample_IgG2	ka (1/Ms)	Stdev	kd (1/s)	Stdev	KD (M)	Stdev	Chi2	Stdev	Sample_IgG3_m15	ka (1/Ms)	Stdev	kd (1/s)	Stdev	KD (M)	Stdev	Chi2	Stdev
Mono					3.06E-08		0.0249		Mono					3.02E-08		0.00957	
3 nm	1.16E+07	2053785.123	3.49E-04	5.7851 E-05	3.08E-11	7.47022 E-12	8.95E-02	0.01114	3 nm	1.10E+07	1584897.473	2.26E-04	2.73632 E-05	2.07E-11	1.93415 E-12	9.48E-02	0.010418
7 nm	1.66E+07	1857057.581	4.53E-04	3.55543 E-05	2.75E-11	3.95875 E-12	4.63E-02	0.004752	7 nm	1.81E+07	1848494.883	1.72E-04	2.25542 E-05	9.44E-12	2.77998 E-13	4.60E-02	0.009655
14 nm	2.79E+07	556497.393	2.46E-04	5.9889 E-05	8.74E-12	6.31372 E-13	3.72E-02	0.01365	14 nm	2.84E+07	4438426.899	2.07E-04	1.29017 E-05	7.38E-12	1.07936 E-12	4.43E-02	0.013424
16 nm	4.18E+07	7750215.051	2.17E-04	1.97308 E-05	5.39E-12	1.63668 E-12	2.77E-02	0.003723	16 nm	4.46E+07	8488221.251	1.10E-04	1.84515 E-05	2.50E-12	5.4079 E-13	3.53E-02	0.006554
17 nm	6.49E+06	130950.3723	3.74E-04	5.67047 E-05	5.76E-11	7.62154 E-12	9.93E-02	0.015001	17 nm	4.62E+06	285310.9415	2.89E-04	6.68755 E-06	6.26E-11	2.66434 E-12	1.42E-01	0.029547
28 nm							1.06E-08	2.33E-02	28 nm	2.90E+05				8.52E-10		4.54E-01	
45 nm							4.03E-08	3.22E-02	45 nm					1.94E-08		8.20E-02	
Sample_IgG3									Sample_IgM_mono								
Sample_IgG3	ka (1/Ms)	Stdev	kd (1/s)	Stdev	KD (M)	Stdev	Chi2	Stdev	Sample_IgM_mono	ka (1/Ms)	Stdev	kd (1/s)	Stdev	KD (M)	Stdev	Chi2	Stdev
Mono					3.36E-08		0.0186		Mono								
3 nm	1.93E+07	1081911.888	2.07E-04	1.46405 E-05	1.07E-11	1.11284 E-12	3.84E-02	0.003612	3 nm	5.08E+07	17794909.38	2.23E-05	1.94429 E-05	3.66E-13	3.17797 E-13	9.44E-02	0.029111
7 nm	1.82E+07	1016021.653	2.32E-04	2.2546 E-05	1.28E-11	1.81596 E-12	3.62E-02	0.002615	7 nm	7.90E+07	22668912.63	5.14E-05	7.78053 E-06	6.68E-13	1.03309 E-13	3.98E-02	0.000624
14 nm	3.08E+07	1055335.65	1.60E-04	3.853 E-06	5.21E-12	5.70263 E-14	2.94E-02	0.005367	14 nm	9.85E+07	5952041.667	5.16E-05	6.52738 E-06	5.26E-13	8.9751 E-14	5.08E-02	0.007107
16 nm	4.57E+07	2494179.892	6.53E-05	6.4144 E-05	1.38E-12	1.28811 E-12	3.06E-02	0.004484	16 nm	1.09E+08	63534334.29	1.36E-04	0.00012792	2.97E-12	4.39863 E-12	5.33E-02	0.044816
17 nm	7.93E+06	793195.016	2.71E-04	8.65005 E-06	3.44E-11	2.92416 E-12	1.07E-01	0.007686	17 nm	3.45E+07	393234.4508	2.95E-04	9.35788 E-06	8.54E-12	3.69091 E-13	1.24E-01	0.02346
28 nm	2.13E+05		2.17E-04				5.11E-10	3.35E-01	28 nm	3.55E+06				7.25E-11		6.67E-01	
45 nm							4.73E-08	1.38E-01	45 nm					7.39E-09		1.47E-01	
Sample_IgG4									Sample_IgM								
Sample_IgG4	ka (1/Ms)	Stdev	kd (1/s)	Stdev	KD (M)	Stdev	Chi2	Stdev	Sample_IgM	ka (1/Ms)	Stdev	kd (1/s)	Stdev	KD (M)	Stdev	Chi2	Stdev
Mono					2.78E-08		0.0158		Mono					3.37E-09		1.88	
3 nm	1.60E+07	1381557.093	4.50E-04	7.37801 E-05	2.81E-11	3.08757 E-12	6.83E-02	0.002875	3 nm	5.29E+07	2806819.077	2.29E-05	1.552 E-06	4.34E-13	5.14413 E-14	5.88E-02	0.00617
7 nm	1.91E+07	994664.9352	2.12E-04	8.88463 E-06	1.11E-11	7.10516 E-13	4.44E-02	0.009283	7 nm	5.08E+07	2751187.622	6.73E-05	1.61327 E-06	6.39E-12	8.95898 E-14	6.36E-02	0.01497
14 nm	3.25E+07	2191483.516	2.09E-04	2.34557 E-05	6.41E-12	3.23153 E-13	5.07E-02	0.013755	14 nm	8.52E+07	4319942.129	2.80E-05	1.98283 E-06	3.29E-13	3.07038 E-14	2.28E-02	0.013271
16 nm	5.08E+07	4318196.962	6.66E-05	1.65004 E-05	1.31E-12	3.10815 E-13	4.81E-02	0.009986	16 nm	1.34E+08	11662046.7	1.98E-05	9.61405 E-07	1.49E-13	1.6129 E-14	4.03E-02	0.004732
17 nm	4.37E+06	572775.9888	3.02E-04	1.28176 E-05	6.97E-11	6.95094 E-12	9.83E-02	0.021984	17 nm	7.19E+07	3150703.625	2.42E-04	1.02632 E-06	3.37E-12	1.5004 E-13	3.57E-02	0.010307
28 nm							5.89E-09	6.22E-02	28 nm					2.42E-04			
45 nm							3.22E-08	2.99E-02	45 nm	4.52E+05				6.64E-04		6.21E+00	
Sample_IgG3_HM4																	
Sample_IgG3_HM4	ka (1/Ms)	Stdev	kd (1/s)	Stdev	KD (M)	Stdev	Chi2	Stdev									
Mono					3.49E-08		0.0246										
3 nm	7.90E+06	152454.3648	8.42E-05	1.10133 E-05	1.07E-11	1.25942 E-12	8.36E-02	0.007536									
7 nm	1.43E+07	830722.577	3.36E-04	1.83058 E-05	2.36E-11	5.48361 E-13	5.06E-02	0.001914									
14 nm	3.27E+07	3016112.288	6.60E-05	1.24018 E-05	2.04E-12	4.79E-13	3.33E-02	0.005173									
16 nm	4.73E+07	5113863.51	3.59E-05	4.53752 E-06	7.61E-13	9.60487 E-14	3.10E-02	0.005006									
17 nm	8.09E+06	168054.5546	4.90E-04	1.1238 E-05	6.06E-11	2.62348 E-12	8.78E-02	0.008793									
28 nm	3.42E+05		4.69E-04				1.22E-09	4.05E-01									
45 nm							3.23E-08	2.36E-02									

Supplementary Table 5. Raw data for the brick digoxigenin hexagon nanopatterns binding to rabbit anti- digoxigenin IgG1. (n = 3)

Sample	ka (1/Ms)	Stdev	kd (1/s)	Stdev	KD (M)	Stdev	Chi2	Stdev
JR1	1.39E+07	683910.3255	6.54E-04	7.12736E-05	4.73E-11	7.67E-12	2.43E-02	0.009257
JR2_Close	1.82E+07	1630163.591	1.02E-04	2.73818E-06	5.60E-12	5.25E-13	2.31E-02	0.011679
JR2_Far	9.45E+06	871458.7387	6.08E-04	6.17846E-06	6.48E-11	6.3E-12	2.97E-02	0.007062
JR3_Close	1.31E+07	999549.8987	2.63E-04	7.78717E-06	2.01E-11	9.06E-13	3.17E-02	0.003119
JR3_Far	9.17E+06	731990.6648	5.06E-04	3.26037E-06	5.55E-11	3.93E-12	4.11E-02	0.000208
JR4_Close	1.85E+07	291433.2399	8.57E-05	5.92276E-06	4.62E-12	2.78E-13	2.64E-02	0.010341
JR4_Far	1.62E+07	525388.7457	6.60E-05	1.50609E-06	4.08E-12	2.15E-13	2.51E-02	0.00896
JR5	1.38E+07	327159.4922	1.58E-04	7.76745E-07	1.14E-11	3.07E-13	2.74E-02	0.004997
JR6	1.58E+07	361708.9069	9.36E-05	1.55104E-06	5.93E-12	2.25E-13	3.38E-02	0.004451

