nature chemistry

Article

https://doi.org/10.1038/s41557-023-01345-4

Heteromultivalency enables enhanced detection of nucleic acid mutations

Received: 29 April 2022

Accepted: 15 September 2023

Published online: 26 October 2023

Check for updates

Brendan R. Deal[1], Rong Ma[1], Steven Narum[2], Hiroaki Ogasawara[1], Yuxin Duan[1], James T. Kindt[1] & Khalid Salaita[1]

Detecting genetic mutations such as single nucleotide polymorphisms (SNPs) is necessary to prescribe effective cancer therapies, perform genetic analyses and distinguish similar viral strains. Traditionally, SNP sensing uses short oligonucleotide probes that differentially bind the SNP and wild-type targets. However, DNA hybridization-based techniques require precise tuning of the probe's binding affinity to manage the inherent trade-off between specificity and sensitivity. As conventional hybridization offers limited control over binding affinity, here we generate heteromultivalent DNA-functionalized particles and demonstrate optimized hybridization specificity for targets containing one or two mutations. By investigating the role of oligo lengths, spacer lengths and binding orientation, we reveal that heteromultivalent hybridization enables fine-tuned specificity for a single SNP and dramatic enhancements in specificity for two non-proximal SNPs empowered by highly cooperative binding. Capitalizing on these abilities, we demonstrate straightforward discrimination between heterozygous *cis* and *trans* mutations and between different strains of the SARS-CoV-2 virus. Our findings indicate that heteromultivalent hybridization offers substantial improvements over conventional monovalent hybridization-based methods.

Specific hybridization between complementary nucleic acids enables many sensing and diagnostic methods[1–4]. For example, polymerase chain reaction (PCR) assays rely on specific hybridization between primers and templates. However, there is often a fundamental trade-off between maximizing specificity, also referred to as the false positive rate in clinical diagnostics, and sensitivity, or true positive rate[5]. High binding affinity results in improved sensitivity, the lowest detectable oligonucleotide concentration, but also leads to enhanced off-target binding and decreased discrimination between similar targets. Conversely, lowering the target affinity can enhance specificity for the complementary target, but lowers the limit of detection of an assay. Thus, there is an affinity 'sweet spot' that maximizes the ratio between on- and off-target binding[6]. Realizing this optimal affinity is difficult, often resulting in poor discrimination for targets containing mismatches, such as single nucleotide polymorphisms (SNPs), which are biomedically relevant and challenging to identify[5,7,8]. One common way to tune affinity to maximize specificity is by changing the probe length. However, the problem with this strategy is that adding or removing a single base pair drastically changes the affinity, resulting in low-precision affinity tuning[6,9]. Adjusting the temperature and ionic strength can precisely optimize the probe affinity for SNP targets, but this approach fails when detecting multiple SNPs simultaneously in a multiplexed or microarray-type assay[10]. Therefore, a fundamental problem in the field pertains to developing facile strategies to fine-tune the target affinity and enhance the specificity.

To overcome this challenge, we tested the hypothesis that multivalent binding can be used to optimize the specificity of hybridization and hence boost the performance of nucleic-acid sensing assays. In many assays, target binding occurs on DNA-functionalized surfaces or particles to allow a more rapid and simple readout[11–14]. These DNA-coated structures, which we refer to as homomultivalent (homoMV; Fig. 1a, top), typically hybridize 'monovalently', forming a

[1]Department of Chemistry, Emory University, Atlanta, GA, USA. [2]Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, GA, USA. e-mail: k.salaita@emory.edu

https://doi.org/10.1038/s41557-023-01345-4

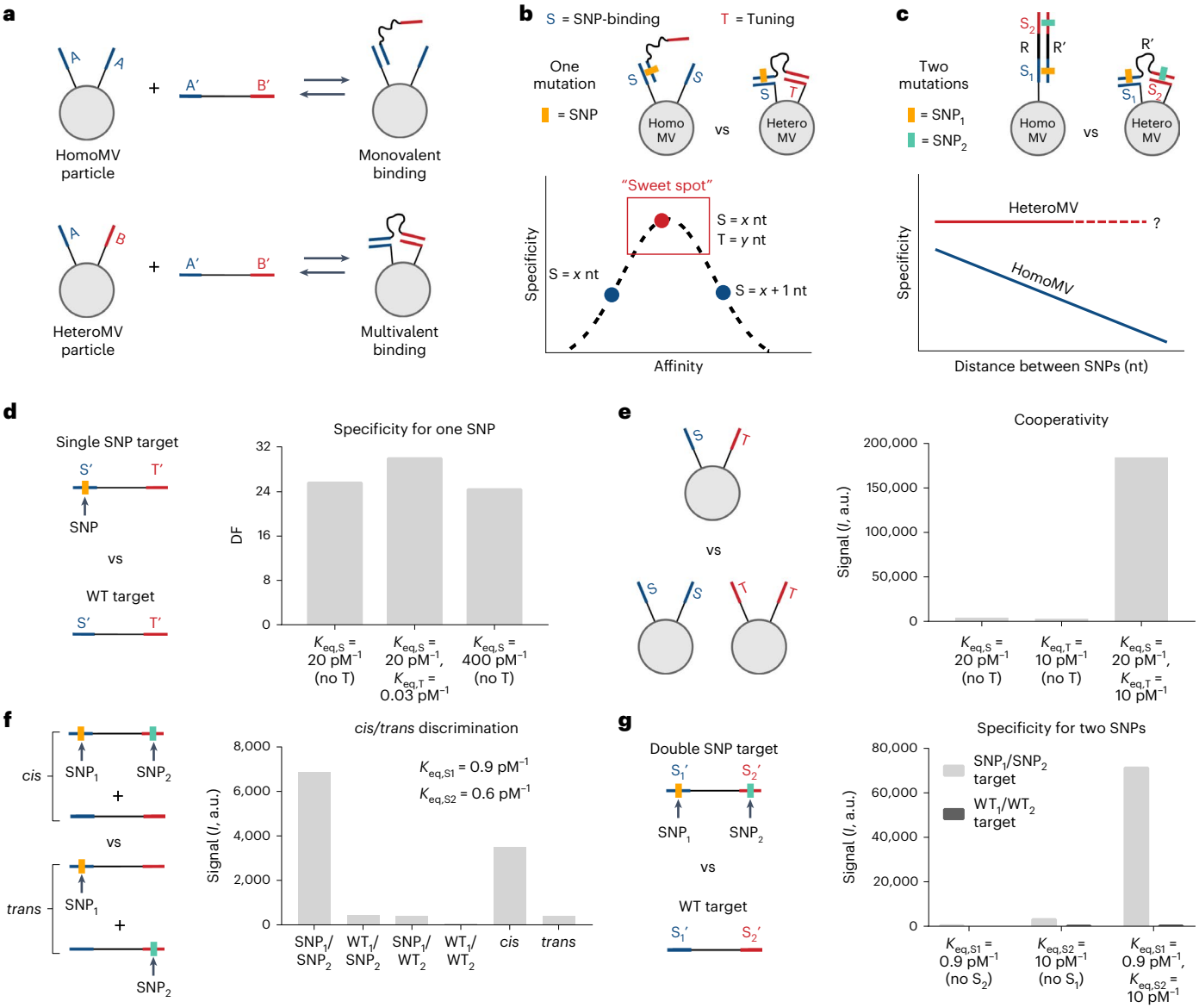


Fig. 1 | Hypothesized advantages of heteromultivalent hybridization and modelling key applications. a, General illustration of a homoMV DNA-coated structure containing only one unique oligonucleotide sequence, A, and a heteroMV DNA-coated structure containing two unique oligonucleotide sequences, A and B. **b**, Scheme illustrating the difficulty in tuning the binding affinity by adding an additional base pair to a homoMV binding interaction and the hypothesized ability of a heteroMV structure to more precisely tune the binding affinity of hybridization to achieve maximum specificity. **c**, Scheme illustrating the hypothesized effect of the distance between two SNPs on homoMV and heteroMV hybridization specificity. **d**–**g**, Schemes and modelling predictions describing the specificity for one SNP (**d**), cooperativity (**e**), *cis*/*trans* discrimination (**f**) and specificity for two SNPs (**g**) of homoMV and heteroMV particles presenting oligos with the K_{eq} values provided.

single duplex with each target. There are a few examples of homoMV structures binding targets multivalently; however, this approach is only applicable for repetitive targets[15,16]. We recently demonstrated that heteromultivalent (heteroMV) structures presenting multiple distinct oligonucleotide sequences (Fig. 1a, bottom) can bind multivalently to non-repetitive targets with high avidity[17]. Here, motivated by this past work, we investigated whether presenting a tuning oligo (T) alongside a SNP-binding oligo (S) can precisely tune the target binding affinity and achieve high specificity for a SNP without relying on buffer optimization (Fig. 1b).

Specificity is also important in applications that require detecting multiple mutations in a single target. For example, haplotype phasing analyses involve distinguishing *cis* and *trans* mutations located on the same or different chromosome copy[18,19]. Differentiating viral strains also requires optimizing specificity for unique mutations. However, detecting two mutations on a target is difficult to achieve, as monovalent binding probes bind either both sites and the region in between (R′) with low specificity (Fig. 1c), or bind each mutation separately with no cooperativity. To address this challenge, we engineered heteroMV binding to hybridize cooperatively to two mutations with a non-complementary spacer in between (Fig. 1c). With heteroMV binding, overall affinity for a desired target is enhanced while maintaining low affinity for single mutant or wild-type targets, similar to 'AND' logic gates and proximity assays[20–22]. Moreover, due to the additive effect of each mismatch, we hypothesized that specificity substantially increases when two mutations are targeted through heteroMV binding. Overall, we aimed to demonstrate that heteroMV binding greatly expands the potential of DNA hybridization-based

assays and DNA nanotechnology by offering highly tunable specificity and cooperativity.

Results

Modelling heteroMV hybridization specificity and cooperativity

To predict the impact of heteroMV binding on hybridization specificity and cooperativity, we focused on an $n = 2$ particle (n corresponds to the number of unique oligosequences anchored to the particle) modified with 50% S and 50% T oligos. Binding of the target to the particle was modelled as a two-step reversible reaction, where S and T bind their complements with association binding constants $K_{eq,S}$ and $K_{eq,T}$, respectively. The particle–target complex can form three distinct binding states where only S binds, only T binds or both segments bind (Extended Data Fig. 1). The equilibrium constant for the target bound to both segments can be described as

$$K_{eq} = K_{eq,S} \times K_{eq,T} \times c_{eff} \quad (1)$$

where c_{eff} is the effective concentration of the unbound second oligo within the volume accessible to the target after binding the first oligo, as described previously[23–26]. Thus, the total affinity of all three states for the particle binding a complementary SNP-containing target is (see the Modelling section of the Methods for further justification)

$$K_{eq,S+T,SNP} = K_{eq,S} + K_{eq,T} + K_{eq,S} \times K_{eq,T} \times c_{eff} \quad (2)$$

To incorporate specificity into the model, we also derived a binding constant for a wild-type target (WT) containing a mismatch in S′ (Fig. 1d, scheme). To account for the decreased affinity of the mismatched S oligo–WT target duplex, a mismatch factor (MM) is multiplied by each $K_{eq,S}$ term in equation (2). Therefore, the total binding affinity for the particle binding the WT target is given by

$$K_{eq,S+T,WT} = MM \times K_{eq,S} + K_{eq,T} + MM \times K_{eq,S} \times K_{eq,T} \times c_{eff} \quad (3)$$

We next derived an equation to calculate the equilibrium binding occupancy, Θ, of the oligos immobilized on the particle surface (see Modelling section of the Methods) and converted Θ to an arbitrary assay signal, I, using inputted maximum and background assay signals with the equation

$$I = I_{max} \times \Theta + I_{bg} \quad (4)$$

By calculating I when the particles bound the SNP target or the WT target, the discrimination factor (DF), a common metric for specificity[6,27], was calculated using the equation

$$DF = I_{SNP}/I_{WT} \quad (5)$$

Moreover, by calculating I when a particle with only the S oligo (I_S), only the T oligo (I_T) or both oligos (I_{S+T}) bound the SNP target (Fig. 1e), the cooperativity factor (CF) was calculated using the following equation to quantify the enhancement in binding due to heteromultivalency:

$$CF = 2 \times I_{S+T}/(I_S + I_T) \quad (6)$$

To predict the impact of $K_{eq,S}$ and $K_{eq,T}$ on DF and CF, we ran numerical analyses using a series of affinities for each oligo with values spanning up to eight orders of magnitude. To roughly approximate the impact of adding one additional base pair to a DNA duplex, the ratio of consecutive values for K_{eq} was chosen to be 20. Mock values of I, DF and CF were then generated for each combination of $K_{eq,S}$ and $K_{eq,T}$, using $c_{eff} = 50$ μM and MM = 0.025 (Extended Data Fig. 1). As described previously for monovalent hybridization[6], the relationship between DF and K_{eq} follows a Gaussian distribution, where a specific K_{eq} value ($K_{eq,optimal}$) maximizes DF (DF_{max}) and any K_{eq} value less than or greater than $K_{eq,optimal}$ results in a diminished DF (Extended Data Fig. 2). For example, our modelling predicts that for an S-only particle, increasing $K_{eq,S}$ from 20 to 400 pM^{-1} (representing the addition of one base pair to the duplex) overshoots $K_{eq,optimal}$ and thus DF_{max} is not achieved. However, adding a T oligo with $K_{eq,T} = 0.03$ pM^{-1}, instead, precisely increases the total affinity from 20 to 40 pM^{-1} and yields a DF greater than that of any of the $n = 1$ particles in the series (Fig. 1d). Note that the model predicts that the T oligo will not enhance DF_{max} and also that if $K_{eq,T}$ is too large (regardless of $K_{eq,S}$), then DF_{max} will decrease (Extended Data Fig. 1).The second major prediction from this simple model is that CF will be greatest when $K_{eq,S} \approx K_{eq,T}$. Specifically, when $K_{eq,S} = 20$ pM^{-1} and $K_{eq,T} = 10$ pM^{-1}, the model predicts that the $n = 2$ particle will bind ~50 times more targets than the average of the two corresponding $n = 1$ particles (Fig. 1e).

We next sought to predict whether heteroMV DNA-coated structures can be used to determine whether two mutations are located on the same or different chromosome copies. Of the ten unique combinations of two mutations on two chromosome copies (Modelling section of the Methods), heterozygous *cis* and *trans* mutations are the most difficult to distinguish (Fig. 1f)[28,29]. To predict the ability to differentiate two *cis* or *trans* mutations, the model was modified so that both oligos are complementary to a SNP (S_1 and S_2 instead of S and T) by applying an MM factor to $K_{eq,S1}$ and $K_{eq,S2}$ when binding a target lacking the corresponding SNPs. This modification then yields equations for total affinity to the SNP_1/SNP_2, SNP_1/WT_2, WT_1/SNP_2 and WT_1/WT_2 targets (Modelling section of the Methods). Equal mixtures of SNP_1/SNP_2 and WT_1/WT_2 targets or SNP_1/WT_2 and WT_1/SNP_2 targets were used to represent heterozygous *cis* or *trans* mutations, respectively. $DF_{cis/trans}$ values were then calculated using

$$DF_{cis/trans} = I_{cis}/I_{trans} \quad (7)$$

Using the same individual oligo binding affinities as used in Fig. 1d,e, the $DF_{cis/trans}$ values were generated for each combination of $K_{eq,S1}$ and $K_{eq,S2}$ (Extended Data Fig. 3). These modelling calculations predicted that two oligos with roughly equal binding affinities, each slightly weaker than those predicted to give the best CF, will result in the highest $DF_{cis/trans}$, yielding a value of 8.4 (Fig. 1f). Alternatively, according to our calculations, to achieve a maximum $DF_{SNP1+SNP2}$ ($I_{SNP1/SNP2}/I_{WT1/WT2}$) value of ~300, the total affinity should be weaker than that which maximizes CF, and stronger than that which maximizes $DF_{cis/trans}$ (Fig. 1g and Extended Data Fig. 3). Note that $DF_{SNP1+SNP2}$ is greatly enhanced due to both binding interactions being impacted by the presence of SNPs. Overall, the mathematical model predicts that a T oligo with lower affinity than the S oligo will give the highest specificity for a single mismatch, a T oligo with similar affinity to the S oligo will maximize cooperativity, and two S oligos with equal but weak affinity will offer the highest *cis*/*trans* discrimination or specificity for targets containing two mutations.

Measuring heteroMV hybridization specificity and cooperativity

To test the modelling predictions, we designed five S oligos (7–11 nt long, 7S–11S) and seven T oligos (4–10 nt long, 4T–10T) complementary to a 25-nt region of the *KRAS* genetic sequence that contains the G12C mutation (Fig. 2a and Supplementary Fig. 1). We focused on this target because *KRAS* is an important oncogene and a driver of lung, pancreatic and colorectal cancers when mutated[30]. The G12C mutant target was perfectly complementary to the S and T oligos, whereas the WT target lacking the mutation binds the S oligo with a single base mismatch and the T oligo with no mismatches. Both targets were modified at their 3′ termini with an Atto647N fluorophore (Supplementary Figs. 2 and 3).

https://doi.org/10.1038/s41557-023-01345-4

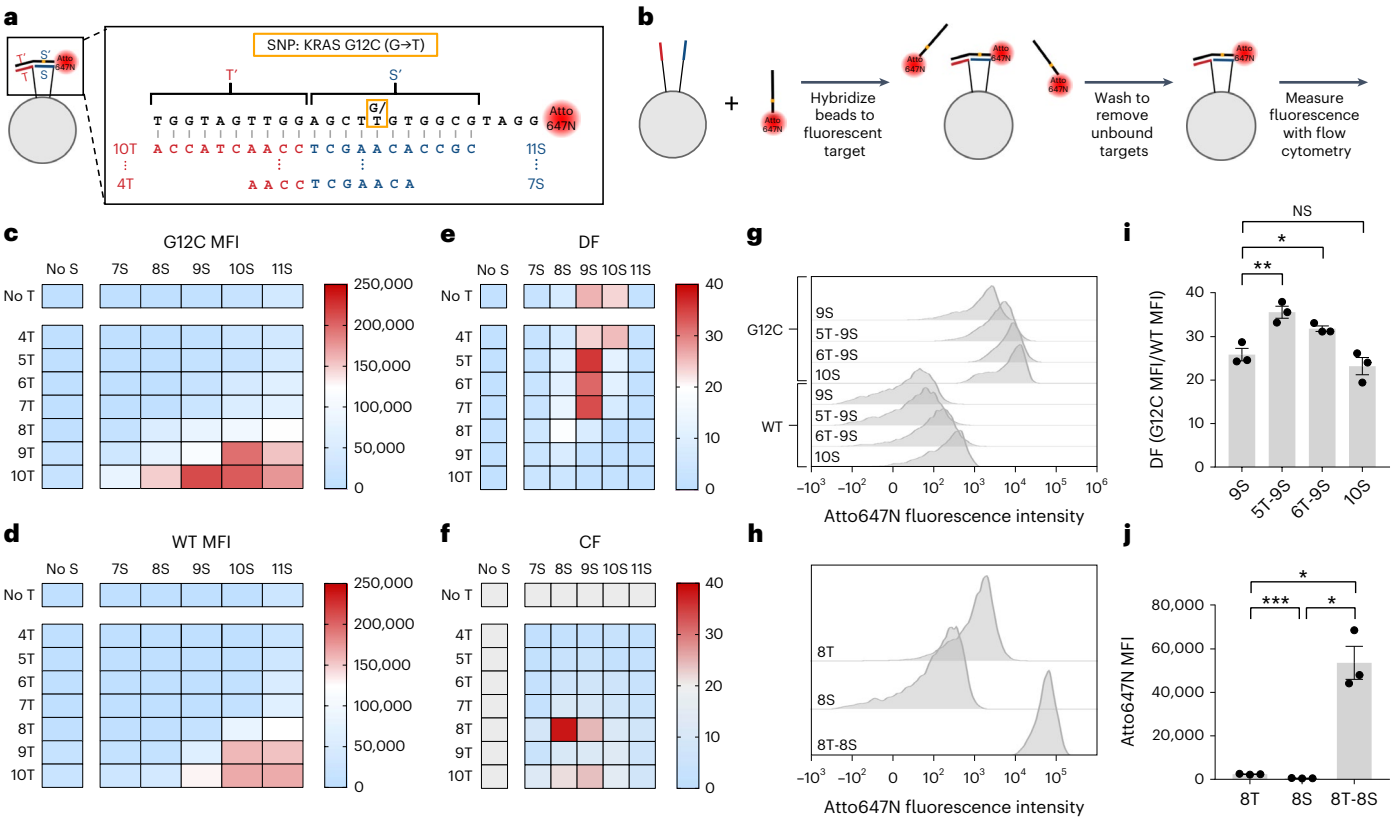


Fig. 2 | Measuring the specificity and cooperativity of heteromultivalent hybridization. a, Design of the oligonucleotides included in the screen to maximize the DF and CF. Yellow boxes describe the SNP and its position in the target sequence. **b**, Scheme describing the flow cytometry-based assay used to quantify target binding to 5-µm DNA-coated silica particles. **c**,**d**, Heatmaps showing the median fluorescence intensity of each bead included in the screen when incubated with the G12C target (**c**) and the WT target (**d**). **e**,**f**, Heatmaps showing the DF (**e**) and CF (**f**) of each bead included in the screen. The CF is shown for beads incubated with the G12C target. **g**, Representative histograms for 9S, 5T–9S, 6T–9S and 10S beads binding the G12C and WT targets. **h**, Representative histograms for 8T, 8S and 8T–8S beads binding the G12C target. **i**, Measured DFs for 9S, 5T–9S, 6T–9S and 10S beads (*P* values: 9S versus 5T–9S = 0.0088, 9S versus 6T–9S = 0.0329, 9S versus 10S = 0.2995). **j**, Measured median fluorescence intensity values for 8T, 8S and 8T–8S beads binding the G12C target (*P* values: 8T versus 8S = 0.0001, 8T versus 8T–8S = 0.0397, 8S versus 8T–8S = 0.0367). Error bars represent the standard error of the mean from $n = 3$ distinct samples. Values were compared using paired one-way analysis of variance (ANOVA) with multiple comparisons follow-up tests ($^{NS}P > 0.05$, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$; NS, not significant).

Each of the S and T oligos contained a T10 polynucleotide linker and a 5′ thiol group to enable conjugation to silica beads (Supplementary Fig. 4). Beads were modified with each possible combination of the S and T oligos, generating a library of 48 unique DNA-coated silica beads. The density of the oligos on the beads was measured to be ~4.1×10^4 oligos µm^{-2} with an average oligo spacing of ~5 nm, allowing S and T oligos to bind multivalently to the same target (Extended Data Fig. 4).

We next designed a flow cytometry-based assay to measure the relative binding of targets to each of the 48 beads. In this assay, the DNA-coated beads were incubated with 1 nM of target in 1× saline sodium citrate (SSC) and 0.1% Tween20 buffer, after which unbound targets were removed through centrifugation and the fluorescence intensity of each individual particle was measured using a flow cytometer (Fig. 2b and Supplementary Fig. 5). As expected, the median fluorescence intensities (MFIs) generally increased when the S and/or the T oligo increased in length, confirming that increasing binding affinity results in higher surface occupancy (Θ) (Fig. 2c,d and Extended Data Fig. 5). To quantify the specificity, DF values were calculated for each bead mixture by dividing the G12C and WT MFIs (Fig. 2e). Consistent with the modelling predictions, the beads presenting the 9S oligo alongside the 5T, 6T or 7T oligo had the highest DFs. Specifically, the 5T–9S beads yielded ~37% higher specificity compared to the 9S beads (Fig. 2i), which had the greatest DF of the homoMV beads tested. Importantly, this enhancement was enabled by precise fine-tuning of K_{eq}, as the 5T–9S and 6T–9S beads yielded MFIs between those of the 9S and 10S beads (Fig. 2g). In further agreement with the modelling, the screen showed that the 8T–8S beads bound most cooperatively to the G12C target, with almost 40 times greater target binding than the average of the 8T and 8S $n = 1$ beads (Fig. 2f,h,j).

Impact of spacer length on heteroMV hybridization

Next, to assess the ability of heteroMV beads to bind with high cooperativity to two non-adjacent regions of a target, several spacer-containing targets were designed and tested. Previously, the impact of long, flexible spacers/linkers on multivalent binding avidity has been a controversial topic. Some studies reported that flexibility leads to poor cooperativity due to loss of conformational entropy upon binding[31], while others noted minimal impacts of spacer length on avidity and cooperativity[15,26]. Hence these experiments were designed to test whether hybridization cooperativity and specificity are maintained when the spacer length increases. We therefore introduced a tri-ethylene glycol (short) or a hexa-ethylene glycol (long) modification between the T′ and S′ binding regions (internal) or, as a negative control, at the 5′ terminus of the targets (terminal) (Fig. 3a). Thus, a total of ten targets were tested with the 8T–8S beads using the flow cytometry-based assay.

The results showed that as the internal spacer length increased, more G12C targets bound the beads (Fig. 3b). Inserting a short spacer also enhanced binding to the WT target, although the long spacer did

https://doi.org/10.1038/s41557-023-01345-4

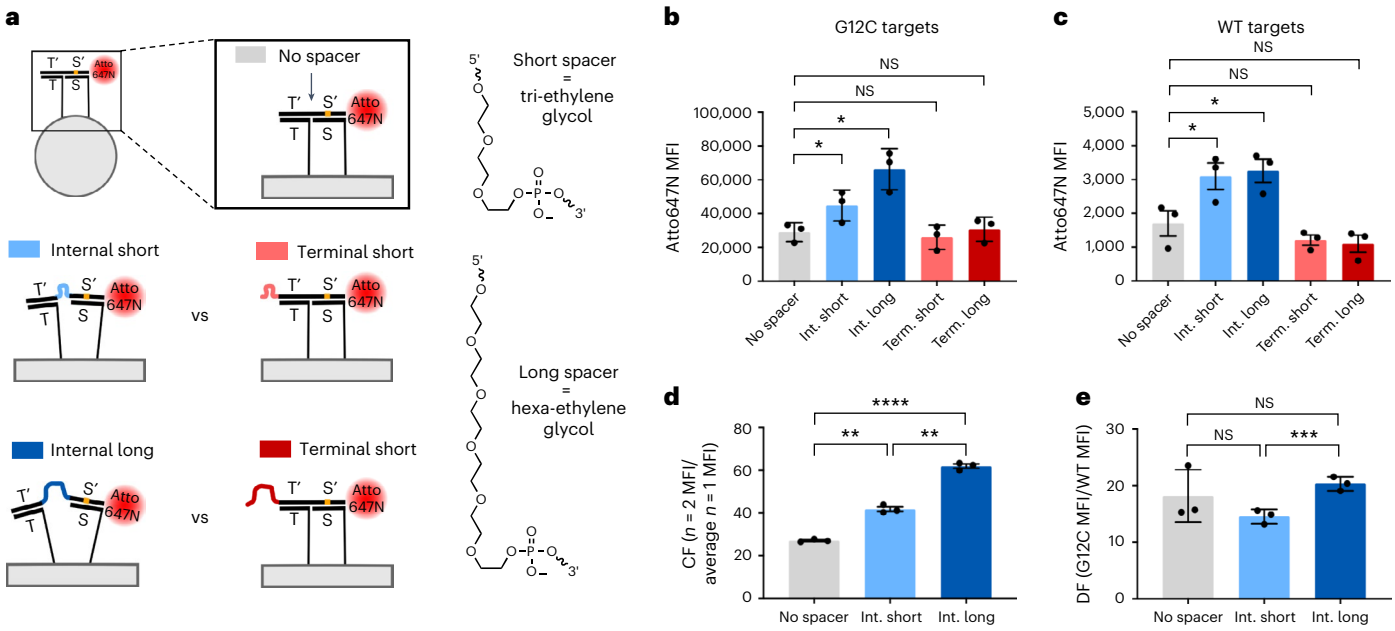


Fig. 3 | Determining the impact of spacer length on heteromultivalent hybridization specificity and cooperativity. a, Scheme describing the design of the no spacer target, the internal and terminal short spacer targets, and the internal and terminal long spacer targets including the chemical structures of the PEG spacer molecules. **b**,**c**, Measured median fluorescence intensity values for 8T–8S beads binding the G12C (**b**; *P* values: no spacer versus internal short = 0.0373, no spacer versus internal long = 0.0222, no spacer versus terminal short = 0.2393, no spacer versus terminal long = 0.4193) and the WT (**c**; *P* values: no spacer versus internal short = 0.0178, no spacer versus internal long = 0.0118, no spacer versus terminal short = 0.3683, no spacer versus terminal long = 0.1138) no spacer, internal short spacer, internal long spacer, terminal short spacer and terminal long spacer targets. **d**,**e**, Measured cooperativity factors (**d**; *P* values: no spacer versus internal short = 0.0068, no spacer versus internal long = <0.0001, internal short versus internal long = 0.0012) and DFs (**e**; *P* values: no spacer versus internal short = 0.6909, no spacer versus internal long = 0.9201, internal short versus internal long = 0.0007) for the 8T–8S beads binding the G12C no spacer, internal short spacer or the internal long spacer targets. Error bars represent the standard error of the mean from $n = 3$ distinct samples. Values were compared using paired one-way ANOVA with multiple comparisons follow-up tests ($^{NS}P > 0.05$, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$). Int., internal; term., terminal.

not lead to a further increase in binding (Fig. 3c and Extended Data Fig. 6). As expected, the terminal spacers did not impact binding to the G12C or WT targets, confirming that the poly-ethylene glycol (PEG) polymer does not chemically influence target binding. The CF of the 8T–8S beads for the G12C targets with different spacer lengths was also calculated by dividing the 8T–8S beads' MFI by the average of the 8T and 8S beads' MFIs when binding the no spacer target. These calculations revealed significant increases in cooperativity as a function of increasing spacer length (Fig. 3d). The impact of spacer length on specificity was also assessed by calculating the DF of the 8T–8S beads for each target. Interestingly, the internal spacers did not lead to a strong effect on specificity relative to the no spacer target, though there was a significant difference in DF between the short and long spacer targets (Fig. 3e). Overall, the investigations into the effect of target spacer length revealed that heteroMV hybridization allows binding to two spacer-separated regions of a target with increased cooperativity and no loss in specificity compared to a target with no spacer. These results will provide guidance in potential designs of proximity or 'AND' logic gate style-assays as well as in diagnostic assays when it is desirable for the tuning oligo to bind a domain (T′) that is not proximal to the SNP site.

Impact of binding orientation on heteroMV hybridization

Due to the antiparallel nature of DNA hybridization, the choice of terminus (5′ or 3′) for the anchoring group of the S and T oligos impacts the direction that the oligo binds the target. Therefore, based on the terminus used for each anchor, the two oligos can bind the target in a head-to-tail, head-to-head or tail-to-tail orientation (Fig. 4a). In this case, head corresponds to the end of the oligo not attached to the particle, and tail corresponds to the linker connecting the oligo to the particle. To understand how binding orientation can potentially impact the properties of the binding interaction, 8T–8S beads that bind in the three different orientations were compared using the flow cytometry-based binding assay. Moreover, to investigate how each orientation is influenced by spacer length, the no spacer, short spacer and long spacer targets were tested with each binding orientation.

When binding the G12C no spacer target, significant differences were observed between the three binding orientations (Fig. 4b,c). Specifically, the head-to-head binding orientation yielded the highest binding, while the tail-to-tail orientation resulted in a greater than threefold reduction in binding compared to the head-to-tail orientation. However, when binding the short or long spacer G12C targets, the tail-to-tail orientation yielded similar binding to the head-to-tail orientation, while the head-to-head orientation still offered slight, non-significant improvements in total binding. Relatedly, the head-to-head orientation beads had a greater than twofold increase in CF relative to the head-to-tail orientation beads and a greater than sixfold increase relative to the tail-to-tail orientation beads when binding the no spacer G12C target (Fig. 4d,e). The greater average CF for the head-to-head orientation was maintained for the spacer-containing targets, although the enhancement was not significant. The results for the WT targets echoed those of the G12C targets (Extended Data Fig. 7). Overall, these results validate the importance of binding orientation in tuning binding affinity and cooperativity.

Together, these results can be explained by considering the effects of both the spacing between segments on the bead surface and the base-stacking interactions at the interface of the T–T′ and S–S′ duplexes. Based on the distance between the T and S oligos on the surface, different binding orientations can minimize energetic strain during binding depending on the linker length and duplex length.

https://doi.org/10.1038/s41557-023-01345-4

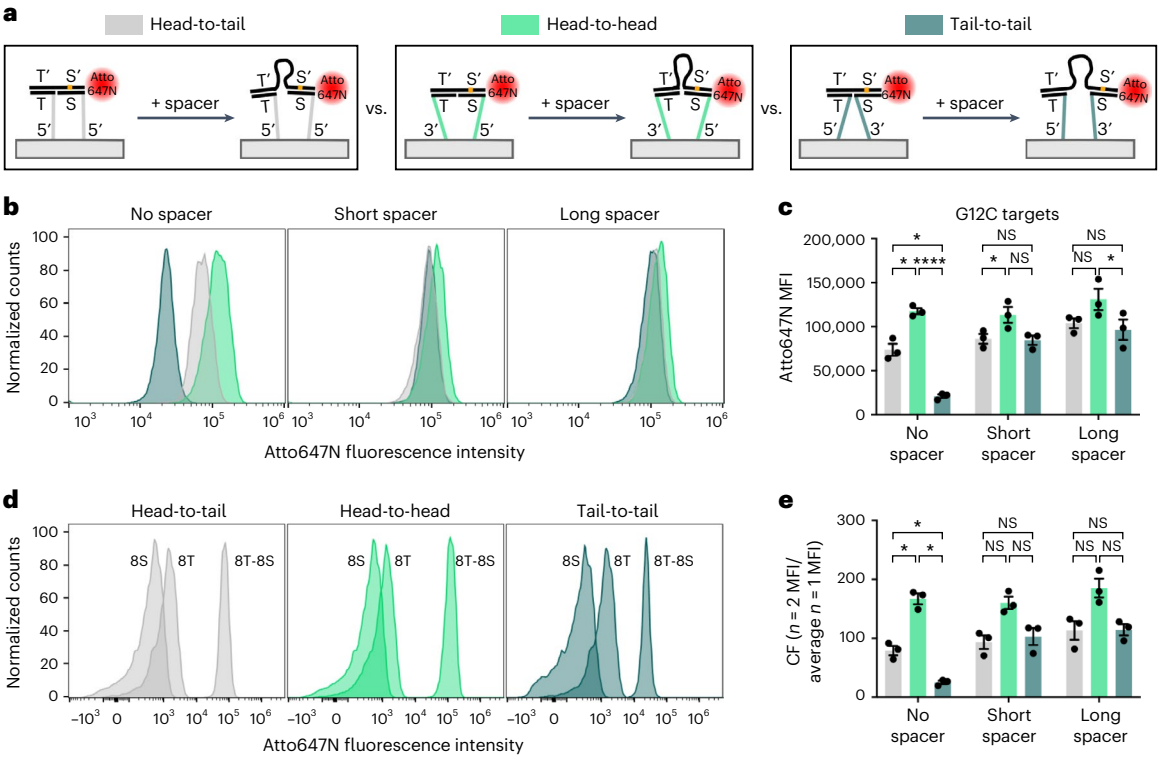


Fig. 4 | Determining the impact of binding orientation on heteromultivalent hybridization specificity and cooperativity. a, Scheme describing $n = 2$ beads with head-to-tail, head-to-head or tail-to-tail orientation binding to targets with or without a spacer region. **b**,**c**, Representative histograms (**b**) and measured median fluorescence intensity values (**c**) for 8T–8S beads with each orientation binding the G12C no spacer, short spacer and long spacer targets (*P* values for no spacer: head-to-tail versus head-to-head = 0.0103, head-to-tail versus tail-to-tail = 0.0177, head-to-head versus tail-to-tail < 0.0001; *P* values for short spacer: head-to-tail versus head-to-head = 0.0266, head-to-tail versus tail-to-tail = 0.8476, head-to-head versus tail-to-tail = 0.0606; *P* values for long spacer: head-to-tail versus head-to-head = 0.148, head-to-tail versus tail-to-tail = 0.588, head-to-head versus tail-to-tail = 0.0199). **d**, Representative histograms for 8T, 8S and 8T–8S beads with each orientation binding the G12C no spacer target. **e**, Measured cooperativity factors for 8T–8S beads with each orientation binding the G12C no spacer, short spacer and long spacer targets (*P* values for no spacer: head-to-tail versus head-to-head = 0.033, head-to-tail versus tail-to-tail = 0.0213, head-to-head versus tail-to-tail = 0.0108; *P* values for short spacer: head-to-tail versus head-to-head = 0.1559, head-to-tail versus tail-to-tail = 0.2322, head-to-head versus tail-to-tail = 0.2492; *P* values for long spacer: head-to-tail versus head-to-head = 0.2515, head-to-tail versus tail-to-tail = 0.9886, head-to-head versus tail-to-tail = 0.1912). Error bars represent the standard error of the mean from $n = 3$ distinct samples. Values were compared using paired one-way ANOVA with multiple comparisons follow-up tests ($^{NS}P > 0.05$, $^{*}P < 0.05$, $^{****}P < 0.0001$).

For example, if T and S are far apart, then binding the no spacer target in the tail-to-tail orientation might result in considerable strain on the T10 linkers. Moreover, previous studies showed that base stacking at a nick site can result in strong enthalpic contributions to overall binding stability[32–35]. This is consistent with the head-to-head orientation yielding the most avid binding as it binds with only a nick between the two duplexes. In contrast, in the other orientations, the T10 linkers probably interfere with this base-stacking interaction and hence reduce binding affinity and cooperativity.

Detecting the *cis*/*trans* relationship of two mutations

We next tested the modelling prediction that heteroMV binding can be used to distinguish heterozygous *cis* and *trans* mutations (Figs. 1f and 5a). This challenging task is critical in medical diagnostics, as the presence of two mutations on the same gene copy can alter protein function, while one mutation on each gene copy can yield cells with no functional gene copies[18,28,29]. Moreover, *cis*/*trans* discrimination is valuable in genetic counselling to track the inheritance of mutations[18]. As a proof of concept, 8- and 9-nt S_1 and S_2 oligos were designed to hybridize in the head-to-tail or head-to-head orientation to a complementary 31-nt target corresponding to a region of the *KRAS* gene that contains the G12C mutation (SNP_1) in the S_1' region and the L19F mutation (SNP_2) in the S_2' region (Fig. 5b and Supplementary Fig. 1). Between the S_1' and S_2' regions there are 13–15 non-complementary nucleotides (Extended Data Fig. 8). L19F is a non-canonical mutation that has been found to cause increased tumour proliferation and transforming potential over WT *KRAS*[36]. We chose to use this mutation in our assay due to its proximity to the G12C mutation (23 nt away), although we anticipate that binding two mutations that are further apart will still be effective.

Using each combination of the binding oligos, eight heteroMV beads were synthesized and flow cytometry was used to measure their binding to 1 nM of the four targets, as well as to a 0.5 nM SNP_1/SNP_2 + 0.5 nM WT_1/WT_2 target mixture (*cis*) or a 0.5 nM SNP_1/WT_2 + 0.5 nM WT_1/SNP_2 target mixture (*trans*) (Fig. 5c,d and Extended Data Fig. 8). As expected, all the bead combinations bound the SNP_1/SNP_2 target with the greatest affinity and the WT_1/WT_2 target with the weakest affinity. Moreover, the $9S_1$–$8S_2$ beads with either binding orientation had weak and approximately equal binding to both single mutant targets while showing strong binding to the SNP_1/SNP_2 target, yielding DF values of ~10 for both mutations. Due to this specificity for both mutations and strong binding cooperativity, both the head-to-tail and head-to-head $9S_1$–$8S_2$ beads bound the *cis* target combination significantly more than the *trans* with $DF_{cis/trans}$ values of 4.7 and 8.4, respectively (Fig. 5e–h). Note that the modelling results gave an identical maximum $DF_{cis/trans}$ value of 8.4. Overall, this screen reveals that heteroMV hybridization enables strong discrimination between *cis* and *trans* heterozygous mutations and demonstrates the importance of precisely tuned binding specificity and cooperativity. This result is

https://doi.org/10.1038/s41557-023-01345-4

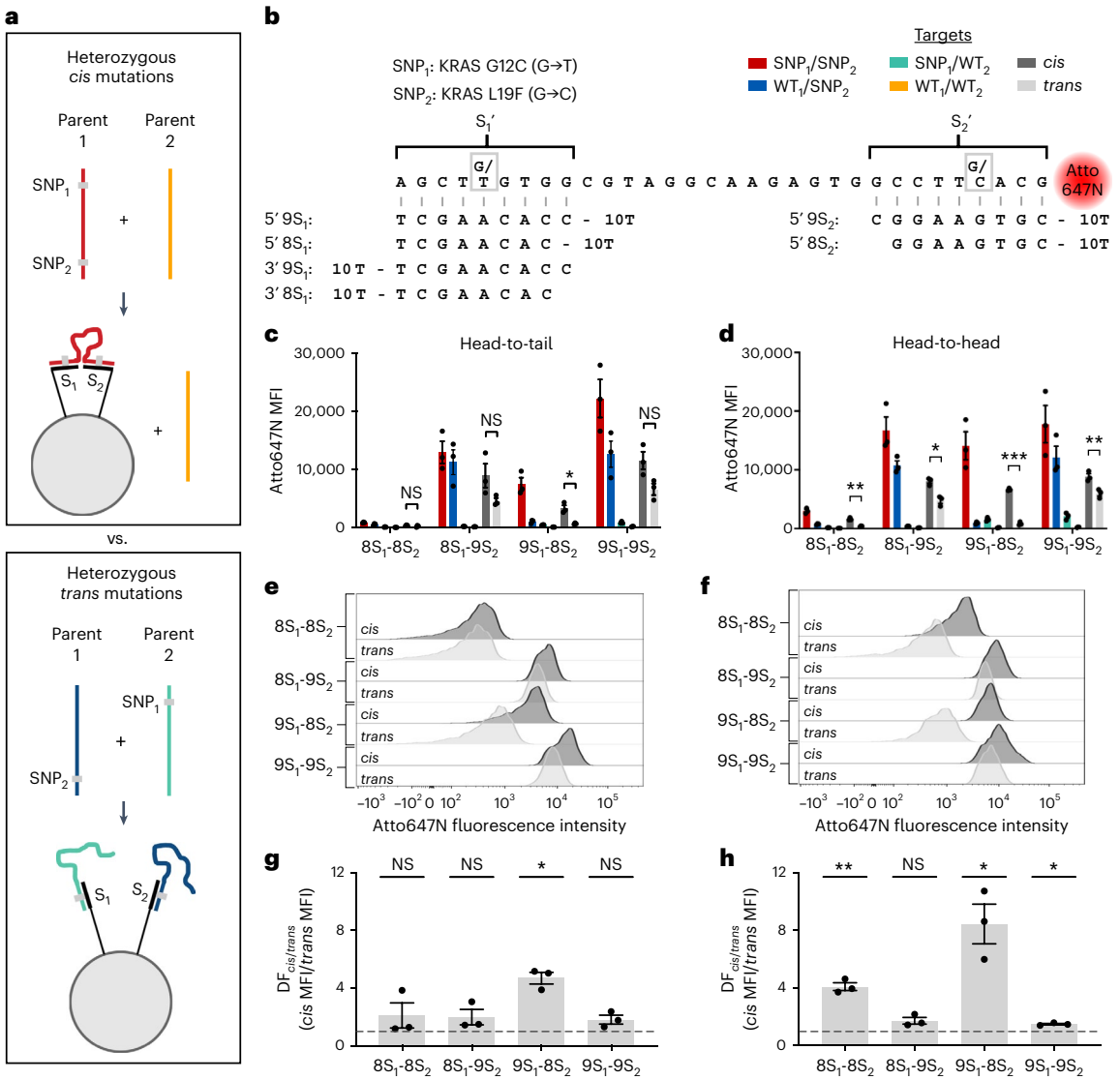


Fig. 5 | Detecting the *cis*/*trans* relationship of two mutations using heteromultivalent hybridization. a, Scheme illustrating the use of head-to-head orientation of heteromultivalent DNA-coated beads to distinguish the heterozygous *cis* mutations mixture (red and yellow targets) from the heterozygous *trans* mutations mixture (blue and green targets). Ideally, the double mutant target (red) will bind the beads multivalently with high affinity, the single mutant targets (blue and green) bind monovalently with low affinity, and the no mutant target (yellow) shows negligible binding. **b**, Scheme describing the sequence of the binding oligos, the identity of the two SNPs, and the two binding orientations tested. **c**,**d**, Measured median fluorescence intensity values for each bead with head-to-tail orientation (**c**; *P* values: $8S_1$–$8S_2$ *cis* versus $8S_1$–$8S_2$ *trans* = 0.1735, $8S_1$–$9S_2$ *cis* versus $8S_1$–$9S_2$ *trans* = 0.1823, $9S_1$–$8S_2$ *cis* versus $9S_1$–$8S_2$ *trans* = 0.0282, $9S_1$–$9S_2$ *cis* versus $9S_1$–$9S_2$ *trans* = 0.0722) or head-to-head orientation (**d**; *P* values: $8S_1$–$8S_2$ *cis* versus $8S_1$–$8S_2$ *trans* = 0.0047, $8S_1$–$9S_2$ *cis* versus $8S_1$–$9S_2$ *trans* = 0.0387, $9S_1$–$8S_2$ *cis* versus $9S_1$–$8S_2$ *trans* = 0.0002, $9S_1$–$9S_2$ *cis* versus $9S_1$–$9S_2$ *trans* = 0.0046) binding each of the targets or target combinations in the legend. Values were compared using two-sided paired Student *t*-tests ($^{NS}P > 0.05$, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$). **e**,**f**, Representative histograms for each bead with head-to-tail orientation (**e**) or head-to-head orientation (**f**) binding the *cis* or *trans* target combinations. **g**,**h**, Measured *cis*/*trans* DFs for each bead with head-to-tail orientation (**g**; *P* values: $8S_1$–$8S_2$ = 0.3318, $8S_1$–$9S_2$ = 0.2029, $9S_1$–$8S_2$ = 0.0119, $9S_1$–$9S_2$ = 0.1169) or head-to-head orientation (**h**; *P* values: $8S_1$–$8S_2$ = 0.0077, $8S_1$–$9S_2$ = 0.0859, $9S_1$–$8S_2$ = 0.0323, $9S_1$–$9S_2$ = 0.0115). Values were compared to a baseline value of 1 (dotted black line) using two-sided one-sample *t*-tests ($^{NS}P > 0.05$, $^{*}P < 0.05$, $^{**}P < 0.01$). Error bars represent the standard error of the mean from $n = 3$ distinct samples.

important as it establishes a hybridization-based approach to distinguish *cis*/*trans* mutations without using enzymes or magnetic separation techniques[29,37–39].

Distinguishing different strains of SARS-CoV-2

We next tested our hypothesis that heteroMV hybridization could lead to dramatic enhancements in specificity for targets containing two mutations (Fig. 1g). We thus designed three model targets corresponding to a 29-nt region of the SARS-CoV-2 spike protein gene that contains three mutations (Q498R, N501Y and Y505H) in the omicron strain, one mutation in the alpha strain (N501Y) and no mutations in the original strain (Fig. 6a). To hybridize specifically to the omicron strain, 8- and 9-nt S_1 and S_2 oligos, complementary to the Q498R site and the Y505H site, respectively, were designed so that neither overlapped with the N501Y mutation shared by the alpha strain (Fig. 6b). Using these oligos, four $n = 2$ beads were synthesized that bound the target in the head-to-head orientation with an 11–13-nt spacer region (Extended Data Fig. 9). As a negative control, $n = 1$ beads functionalized with a 29-nt oligo that is perfectly complementary to the omicron target were also tested (Fig. 6b). Flow cytometry results showed that each of

https://doi.org/10.1038/s41557-023-01345-4

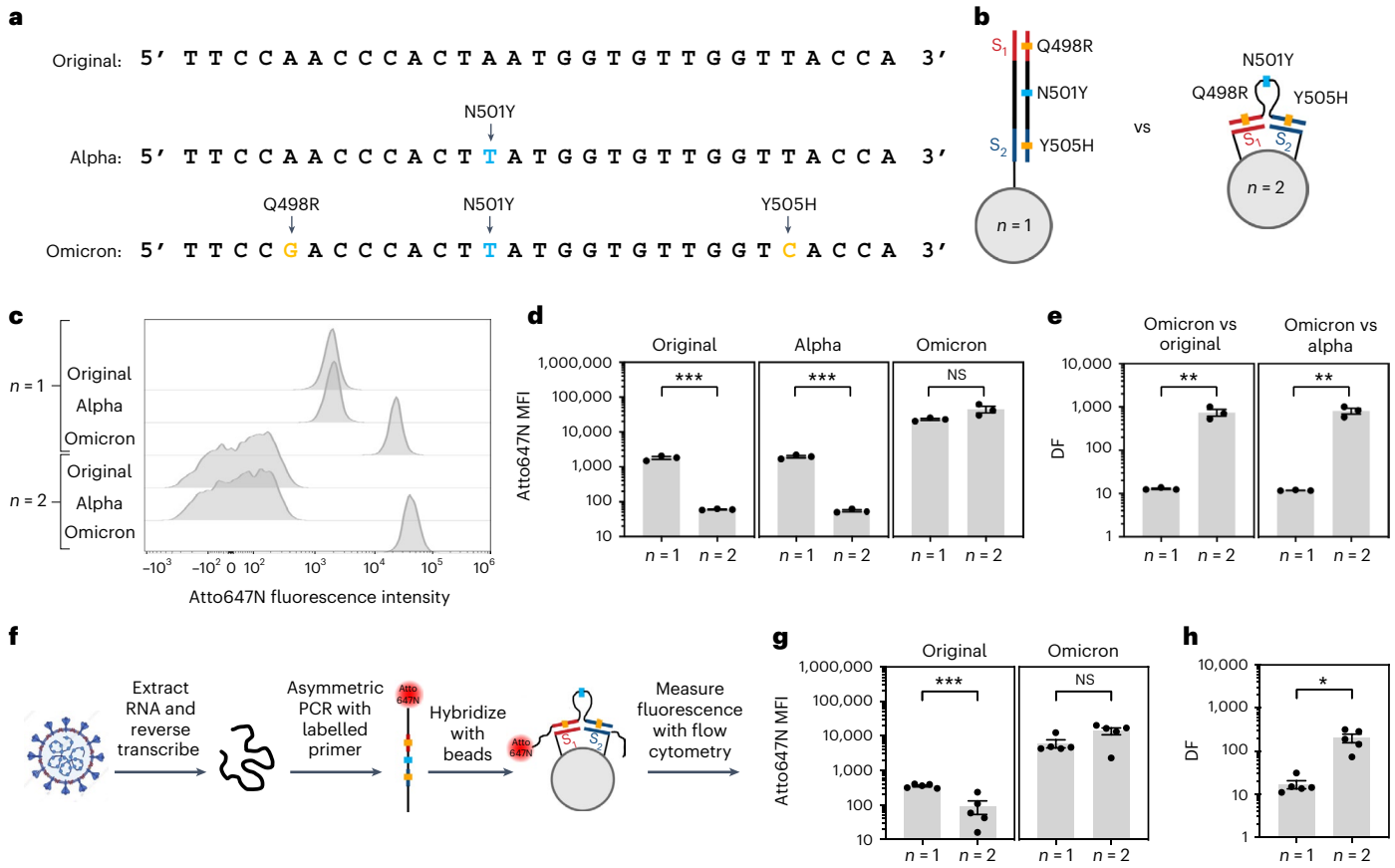


Fig. 6 | Distinguishing different strains of SARS-CoV-2 using heteromultivalent hybridization. a, Sequences of targets based on the original, alpha and omicron strains of the SARS-CoV-2 spike protein, with the mutations in each target indicated with arrows. **b**, Scheme describing the binding of an $n = 1$ bead functionalized with an oligo that is fully complementary to the omicron target and the binding of an $n = 2$ bead functionalized with S_1 and S_2 oligos that are complementary to the regions of the target containing the Q498R and Y505H mutations but not the N501Y mutation. **c**,**d**, Representative histograms (**c**) and measured median fluorescence intensity values (**d**) for the $n = 1$ and $8S_1$–$9S_2$ $n = 2$ beads binding each target (*P* values: original = 0.0005, alpha = 0.0002, omicron = 0.0851). **e**, Measured discrimination factors for the $n = 1$ and $8S_1$–$9S_2$ $n = 2$ beads binding the omicron target versus the original target or the omicron target versus the alpha target (*P* values: omicron versus original = 0.0061, omicron versus alpha = 0.003). **f**, Scheme describing the workflow used to extract RNA from SARS-CoV-2 viral particles, amplify Atto647N-labelled single-stranded targets, and measure binding with flow cytometry. **g**, Measured median fluorescence intensity values for the $n = 1$ and $9S_1$–$9S_2$ $n = 2$ beads binding each virus extracted target (*P* values: original = 0.0006, omicron = 0.0627). **h**, Measured discrimination factors for the $n = 1$ and $9S_1$–$9S_2$ $n = 2$ beads binding the virus extracted omicron target versus the original target ($P = 0.0179$). Values in **d** and **e** were compared using two-sided unpaired Student *t*-tests, and values in **g** and **h** were compared using two-sided paired Student *t*-tests ($^{NS}P > 0.05$, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$). Error bars represent the standard error of the mean from $n = 3$ distinct samples in **d** and **e** and $n = 5$ distinct samples in **g** and **h**.

the $n = 2$ beads tested bound to the omicron target with similarly high affinity and showed minimal binding to the alpha and original targets (Extended Data Fig. 9). Meanwhile, compared to the $n = 2$ beads, the $n = 1$ beads yielded an approximately equal MFI when binding the omicron target but bound to significantly more alpha and original targets (Fig. 6c,d). Importantly, the $n = 2$ beads offered dramatically enhanced specificity for the omicron strain, with the $8S_1$–$9S_2$ combination bead giving a $DF_{SNP1+SNP2}$ value of ~800 compared to either of the other targets (Fig. 6e). The $n = 1$ bead had much lower specificity for the omicron target, with $DF_{SNP1+SNP2}$ values of ~12.

To further demonstrate the utility of heteroMV hybridization for distinguishing viral strains, we compared the ability of $n = 1$ and $n = 2$ beads to discriminate between nucleic acid targets isolated from authentic original strain and omicron strain SARS-CoV-2 viral particles. Following RNA extraction and reverse transcription, complementary DNA (cDNA) was amplified using asymmetric PCR using a tenfold excess of an Atto647N-labelled forward primer (Fig. 6f and Extended Data Fig. 10). This protocol generated fluorescent 88-nt DNA oligos from SARS-CoV-2 virions. Each target (10 nM) was then mixed with the $n = 1$ and $9S_1$–$9S_2$ $n = 2$ beads described above, and binding was analysed using flow cytometry. Note that the $n = 1$ bead binds the original target with three mismatches, whereas the $n = 2$ binds with only two mismatches. Despite this disadvantage, the $n = 2$ bead resulted in a nearly 12-fold increase in DF compared to the $n = 1$ bead as a result of similar binding to the omicron target and reduced binding to the original target (Fig. 6g,h).

As the $n = 1$ bead has more total complementarity with the targets, it was surprising that the $n = 1$ and $n = 2$ beads yielded approximately equal omicron target binding. Potential explanations include increased secondary structure, reduced k_{on} rates and reduced DNA density for the $n = 1$ bead, as has been previously observed for materials functionalized with longer oligos[40–42]. This highlights a general advantage for heteroMV hybridization where each oligo can be shorter in length and therefore less likely to be impacted by these issues. Moreover, the stark differences in specificity between the $n = 1$ and $n = 2$ beads would probably become even greater as the inter-SNP distance increases (Fig. 1c). In this case, the length of the oligo on the $n = 1$ bead would have to become longer to bind to both SNPs, while the oligos on the $n = 2$ beads would not need to be altered, and instead potentially exhibit stronger and more cooperative binding as shown in Fig. 4. This demonstration

of rapid and effective identification of the strain of model viral targets using heteroMV hybridization has the potential to greatly impact the fields of diagnostics, medicine and public health.

Discussion

In this Article we have shown that densely coating a microparticle with two distinct oligonucleotide sequences yields customizable multivalent binding with highly tunable affinity. This result led to several important capabilities. By first optimizing each oligo's length, we have shown that heteroMV binding can control the binding strength more precisely than monovalent binding, enabling near-maximum discrimination of SNPs. Thus, heteroMV hybridization offers an approach to optimizing the performance of hybridization-based mutation detection tools while maintaining compatibility with multiplex assays. Although different mutations and assay conditions will still require optimization of the oligo lengths to tune specificity, the results herein will accelerate future screening processes. Moreover, heteroMV binding can be combined with other approaches that are commonly used to enhance binding specificity, such as molecular beacons, toehold-mediated hybridization and competition/sink probes[3,5–8].

In addition to adjusting the oligo length, customizing the spacer length and binding orientation allowed the demonstration of highly cooperative binding to two unique regions of a target. Both parameters are thus critical for applications that necessitate selective hybridization only when two receptors are present[21]. Enhanced cooperativity was also observed as the spacer length increased, potentially due to an improved ability for a target to reach two adjacent surface oligos. Additionally, as the spacer length increases, the target can span longer distances on the particle surface, allowing access to more copies of each binding oligo. These added binding partners, although spread through a larger volume, can result in a higher local concentration of surface-bound oligos[24]. This feature is unique to heteroMV structures that are densely functionalized, as opposed to a structure that presents a single copy of each oligo and thus cannot access additional binding sites despite a longer spacer. Cooperative binding was demonstrated with up to 15-nt spacers, although further studies with longer spacers would deepen the investigation. Furthermore, when binding the target lacking a spacer, a sixfold increase in cooperativity was observed when head-to-head orientation was used instead of tail-to-tail. However, for heteroMV binding where $n > 2$, it is not possible to exclusively use the highly cooperative head-to-head orientation. Instead, each adjacent oligo pair must alternate between binding in the head-to-head and tail-to-tail orientation or each oligo can be anchored through the same terminus, as previously demonstrated[17].

Through the combined benefits of highly tunable affinity and strong cooperativity despite a spacer region in the target, heteroMV binding also resulted in the ability to distinguish heterozygous *cis* and *trans* mutations. Through optimization, approximately eightfold higher binding was observed when heteroMV particles were incubated with a mixture of double mutant and non-mutant targets rather than two single mutant targets. Distinguishing between these target mixtures is often achieved through costly and lengthy methods involving complex next-generation sequencing assays, droplet PCR or single-molecule dilution[18,19]. Alternatively, in monovalent hybridization-based assays, either one long probe is used to bind both mutations or a distinct probe binds each mutation. In the first case, specificity and cooperativity diminish due to excessively strong binding, and in the second case each probe binds identically to *cis* and *trans* target mixtures[29]. For this reason, hybridization-based assays typically rely on a second discriminatory step involving enzymes or separation techniques[29,37,38]. Finally, heteroMV hybridization enabled ~200-fold higher binding to targets extracted from omicron strain SARS-CoV-2 viral particles compared to original strain particles. Note that whole-genome sequencing is typically performed for strain identification. Thus, the ability to rapidly determine the strain of the viral sample offers facile monitoring of viral evolution.

The heteroMV hybridization approach presented herein is compatible with many materials used to present oligos in close proximity, including one-, two- or three-dimensional structures[15,16]. Also, precisely controlling the inter-oligo distance on the surface is not necessary when target binding regions are further apart, as such targets can span longer distances without diminished cooperativity. The cooperativity arising from heteroMV binding does, however, depend on the oligos being pre-linked to a scaffold. Alternatively, 'binary' probes have been described, which rely on monovalent binding of two unlinked oligo probes and a separate complex formation step to generate a signal[10,43]. Additionally, DNA origami nanoswitches have been engineered to switch to a loop conformation upon heteroMV binding to a target to facilitate detection with gel electrophoresis[22,44]. In this work, fluorophore-labelled targets were used to enable a rapid flow cytometry readout, and an asymmetric PCR approach was demonstrated for diagnostic applications requiring unlabelled target sensing. In many nucleic acid detection methods, tunable binding affinity that allows highly specific and cooperative binding is essential, so heteroMV DNA hybridization is a promising method for further advancing biomedical sensing and diagnostics.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41557-023-01345-4.

References

1. Gunderson, K. L., Steemers, F. J., Lee, G., Mendoza, L. G. & Chee, M. S. A genome-wide scalable SNP genotyping assay using microarray technology. *Nat. Genet.* **37**, 549–554 (2005).
2. Koltai, H. & Weingarten-Baror, C. Specificity of DNA microarray hybridization: characterization, effectors and approaches for data correction. *Nucleic Acids Res.* **36**, 2395–2405 (2008).
3. Tyagi, S., Bratu, D. P. & Kramer, F. R. Multicolor molecular beacons for allele discrimination. *Nat. Biotechnol.* **16**, 49–53 (1998).
4. Tulpan, D. et al. Thermodynamically based DNA strand design. *Nucleic Acids Res.* **33**, 4951–4964 (2005).
5. Chen, X. et al. Thermodynamics and kinetics guided probe design for uniformly sensitive and specific DNA hybridization without optimization. *Nat. Commun.* **10**, 4675 (2019).
6. Zhang, D. Y., Chen, S. X. & Yin, P. Optimizing the specificity of nucleic acid hybridization. *Nat. Chem.* **4**, 208–214 (2012).
7. Wang, J. S. & Zhang, D. Y. Simulation-guided DNA probe design for consistently ultraspecific hybridization. *Nat. Chem.* **7**, 545–553 (2015).
8. Tyagi, S. & Kramer, F. R. Molecular beacons: probes that fluoresce upon hybridization. *Nat. Biotechnol.* **14**, 303–308 (1996).
9. Suzuki, S., Ono, N., Furusawa, C., Kashiwagi, A. & Yomo, T. Experimental optimization of probe length to increase the sequence specificity of high-density oligonucleotide microarrays. *BMC Genomics* **8**, 373 (2007).
10. Kolpashchikov, D. M. Binary probes for nucleic acid analysis. *Chem. Rev.* **110**, 4709–4723 (2010).
11. Taton, T. A., Mirkin, C. A. & Letsinger, R. L. Scanometric DNA array detection with nanoparticle probes. *Science* **289**, 1757–1760 (2000).
12. Alhasan, A. H. et al. Scanometric microRNA array profiling of prostate cancer markers using spherical nucleic acid-gold nanoparticle conjugates. *Anal. Chem.* **84**, 4153–4160 (2012).
13. Diehl, F. et al. Detection and quantification of mutations in the plasma of patients with colorectal tumors. *Proc. Natl Acad. Sci. USA* **102**, 16368–16373 (2005).

14. Schena, M., Shalon, D., Davis, R. W. & Brown, P. O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467–470 (1995).
15. Curk, T. et al. Computational design of probes to detect bacterial genomes by multivalent binding. *Proc. Natl Acad. Sci. USA* **117**, 8719–8726 (2020).
16. Estirado, E. M., Aleman Garcia, M. A., Schill, J. & Brunsveld, L. Multivalent ultrasensitive interfacing of supramolecular 1D nanoplatforms. *J. Am. Chem. Soc.* **141**, 18030–18037 (2019).
17. Deal, B. R. et al. Engineering DNA-functionalized nanostructures to bind nucleic acid targets heteromultivalently with enhanced avidity. *J. Am. Chem. Soc.* **142**, 9653–9660 (2020).
18. Regan, J. F. et al. A rapid molecular approach for chromosomal phasing. *PLoS ONE* **10**, e0118270 (2015).
19. Zheng, G. X. et al. Haplotyping germline and cancer genomes with high-throughput linked-read sequencing. *Nat. Biotechnol.* **34**, 303–311 (2016).
20. Song, T. et al. Fast and compact DNA logic circuits based on single-stranded gates using strand-displacing polymerase. *Nat. Nanotechnol.* **14**, 1075–1081 (2019).
21. Schueder, F. et al. Super-resolution spatial proximity detection with proximity-PAINT. *Angew. Chem. Int. Ed.* **60**, 716–720 (2021).
22. Chandrasekaran, A. R. et al. Cellular microRNA detection with miRacles: microRNA-activated conditional looping of engineered switches. *Sci. Adv.* **5**, eaau9443 (2019).
23. Krishnamurthy, V. M., Semetey, V., Bracher, P. J., Shen, N. & Whitesides, G. M. Dependence of effective molarity on linker length for an intramolecular protein-ligand system. *J. Am. Chem. Soc.* **129**, 1312–1320 (2007).
24. Huskens, J. et al. A model for describing the thermodynamics of multivalent host-guest interactions at interfaces. *J. Am. Chem. Soc.* **126**, 6784–6797 (2004).
25. Sorensen, C. S. & Kjaergaard, M. Effective concentrations enforced by intrinsically disordered linkers are governed by polymer physics. *Proc. Natl Acad. Sci. USA* **116**, 23124–23131 (2019).
26. Kane, R. S. Thermodynamics of multivalent interactions: influence of the linker. *Langmuir* **26**, 8636–8640 (2010).
27. Wu, T., Xiao, X., Zhang, Z. & Zhao, M. Enzyme-mediated single-nucleotide variation detection at room temperature with high discrimination factor. *Chem. Sci.* **6**, 1206–1211 (2015).
28. Chen, N. & Schrijver, I. Allelic discrimination of *cis-trans* relationships by digital polymerase chain reaction: GJB2 (p.V27I/p.E114G) and CFTR (p.R117H/5T). *Genet. Med.* **13**, 1025–1031 (2011).
29. Fan, T. W., Yu, H. L. L. & Hsing, I. M. Conditional displacement hybridization assay for multiple SNP phasing. *Anal. Chem.* **89**, 9961–9966 (2017).
30. Yu, H. A. et al. Prognostic impact of KRAS mutation subtypes in 677 patients with metastatic lung adenocarcinomas. *J. Thorac. Oncol.* **10**, 431–437 (2015).
31. Mammen, M., Choi, S. K. & Whitesides, G. M. Polyvalent interactions in biological systems: implications for design and use of multivalent ligands and inhibitors. *Angew. Chem. Int. Ed.* **37**, 2754–2794 (1998).
32. Yakovchuk, P., Protozanova, E. & Frank-Kamenetskii, M. D. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res.* **34**, 564–574 (2006).
33. Lane, M. J. et al. The thermodynamic advantage of DNA oligonucleotide 'stacking hybridization' reactions: energetics of a DNA nick. *Nucleic Acids Res.* **25**, 611–617 (1997).
34. Maldonado-Rodriguez, R., Espinosa-Lara, M., Loyola-Abitia, P., Beattie, W. G. & Beattie, K. L. Mutation detection by stacking hybridization on genosensor arrays. *Mol. Biotechnol.* **11**, 13–25 (1999).
35. Walter, A. E. et al. Coaxial stacking of helixes enhances binding of oligoribonucleotides and improves predictions of RNA folding. *Proc. Natl Acad. Sci. USA* **91**, 9218–9222 (1994).
36. Munoz-Maldonado, C., Zimmer, Y. & Medova, M. A comparative analysis of individual RAS mutations in cancer biology. *Front. Oncol.* **9**, 1088 (2019).
37. Lee Yu, H. L., Fan, T. W. & Hsing, I. M. Oligonucleotide hybridization with magnetic separation assay for multiple SNP phasing. *Anal. Chim. Acta X* **5**, 100050 (2020).
38. Zhuang, X., Yu, H. L. L. & Hsing, I. M. Toehold probe-based interrogation for haplotype phasing of long nucleic acid strands. *Anal. Methods* **12**, 4185–4190 (2020).
39. Chang, W. et al. Molecular AND logic gate for multiple single-nucleotide mutations detection based on CRISPR/Cas9n system-trigged signal amplification. *Anal. Chim. Acta* **1112**, 46–53 (2020).
40. Gao, Y., Wolf, L. K. & Georgiadis, R. M. Secondary structure effects on DNA hybridization kinetics: a solution versus surface comparison. *Nucleic Acids Res.* **34**, 3370–3377 (2006).
41. Bazrafshan, A. et al. DNA gold nanoparticle motors demonstrate processive motion with bursts of speed up to 50 nm per second. *ACS Nano* **15**, 8427–8438 (2021).
42. Yehl, K. et al. Catalytic deoxyribozyme-modified nanoparticles for RNAi-independent gene regulation. *ACS Nano* **6**, 9150–9157 (2012).
43. Karadeema, R. J., Stancescu, M., Steidl, T. P., Bertot, S. C. & Kolpashchikov, D. M. The owl sensor: a 'fragile' DNA nanostructure for the analysis of single nucleotide variations. *Nanoscale* **10**, 10116–10122 (2018).
44. Zhou, L. et al. Sequence-selective purification of biological RNAs using DNA nanoswitches. *Cell. Rep. Methods* **1**, 100126 (2021).

Methods

Modelling

In all modelling calculations herein, c_{eff} = 50 μM, MM = 0.025, and the ratio of consecutive values of K_{eq} was chosen to be 20. These values were selected to most closely reproduce the obtained experimental results. Note that the equation for total binding affinity for an arbitrary n = 2 bead–target complex ($K_{eq,\,S+T}$, equation (2)) is derived from summing the individual K_{eq} terms ($K_{eq,\,S}$, $K_{eq,\,T}$ and $K_{eq,\,ST}$) of the three possible binding states, as these binding states cannot be distinguished in our experimental assay. To derive an equation to calculate the equilibrium binding occupancy, Θ, of the oligos coating the particle surface using the total binding affinity of the target (K_{eq}, calculated from equations (2) and (3)), the total target concentration ([target]) and the total concentration of oligos on the particle surface ([surface]), we began with the standard equation for K_{eq}:

$$K_{eq} = \frac{[\text{surface}_{bound}]}{[\text{target}_{unbound}][\text{surface}_{unbound}]} \tag{8}$$

Note that K_{eq} represents the binding association constant and thus has units of inverse concentration. This equation is then rearranged into the following form:

$$K_{eq}[\text{target}_{unbound}] = \frac{[\text{surface}_{bound}]}{[\text{surface}_{unbound}]} \tag{9}$$

$[\text{surface}_{bound}]$ is then replaced with $[\text{surface}_{total}] - [\text{surface}_{unbound}]$ to give the following equation:

$$K_{eq}[\text{target}_{unbound}] = \frac{[\text{surface}_{total}] - [\text{surface}_{unbound}]}{[\text{surface}_{unbound}]} \tag{10}$$

As $([\text{surface}_{total}] - [\text{surface}_{unbound}])/[\text{surface}_{total}] = \Theta$ and $[\text{surface}_{unbound}]/[\text{surface}_{total}] = 1 - \Theta$, the following equation is then derived:

$$\frac{\Theta}{1-\Theta} = K_{eq}[\text{target}_{unbound}] \tag{11}$$

To solve for Θ using only K_{eq}, [target] and [surface], $[\text{target}_{unbound}]$ is first replaced with $[\text{target}] - [\text{target}_{bound}]$:

$$\frac{\Theta}{1-\Theta} = K_{eq}([\text{target}] - [\text{target}_{bound}]) \tag{12}$$

$[\text{target}_{bound}]$ is replaced with $[\text{surface}]\Theta$ to give the following equation:

$$\frac{\Theta}{1-\Theta} = K_{eq}([\text{target}] - [\text{surface}]\Theta) \tag{13}$$

This equation is then rearranged into a quadratic form ($\Theta = a\Theta^2 + b\Theta + c$) as follows:

$$\Theta = (1-\Theta)(K_{eq}([\text{target}] - K_{eq}[\text{surface}]\Theta) \tag{14}$$

$$\Theta = K_{eq}[\text{surface}]\Theta^2 - K_{eq}[\text{target}]\Theta - K_{eq}[\text{surface}]\Theta + K_{eq}[\text{target}] \tag{15}$$

$$0 = K_{eq}[\text{surface}]\Theta^2 - K_{eq}[\text{target}]\Theta - K_{eq}[\text{surface}]\Theta - \Theta + K_{eq}[\text{target}] \tag{16}$$

$$0 = K_{eq}[\text{surface}]\Theta^2 + (-K_{eq}([\text{target}] - [\text{surface}]) - 1)\Theta + K_{eq}[\text{target}] \tag{17}$$

The equation is then solved for Θ using the quadratic formula

$$\Theta = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{18}$$

where $a = K_{eq}[\text{surface}]$, $b = -K_{eq}([\text{target}] - [\text{surface}]) - 1$, and $c = K_{eq}[\text{target}]$, giving the final equation for Θ:

$$\Theta = \frac{K_{eq}([\text{target}] + [\text{surface}]) + 1 - \sqrt{(K_{eq}([\text{target}] + [\text{surface}]) + 1)^2 - 4K_{eq}^2[\text{surface}][\text{target}]}}{2K_{eq}[\text{surface}]} \tag{19}$$

Note that the correct value of Θ is equal to the root given by subtracting the quadratic portion, and thus the ± term was replaced with a − term in equation (19) and the root given by adding the quadratic portion was ignored. For the results shown in Fig. 1d,e,g, [target] = 1 nM and [surface] = 1 nM, and thus the following simplifications can be made, where $K = K_{eq} \times 1\,\text{nM}$:

$$\Theta = \frac{2K + 1 - \sqrt{(2K+1)^2 - 4K^2}}{2K} \tag{20}$$

$$\Theta = \frac{1}{2K} + 1 - \frac{\sqrt{4K^2 + 4K + 1 - 4K^2}}{2K} \tag{21}$$

$$\Theta = \frac{1}{2K} + 1 - \frac{\sqrt{4K+1}}{2K} \tag{22}$$

For the results shown in Fig. 1f, $[\text{target}_1]$ = 0.5 nM and $[\text{target}_2]$ = 0.5 nM. Therefore, the values of Θ for target_1 and target_2 are calculated using Microsoft Excel and then summed to calculate the total Θ. Finally, Θ was converted to an arbitrary assay signal, I, using the equation

$$I = I_{max} \times \Theta + I_{bg} \tag{23}$$

where I_{max} represents the maximum signal and I_{bg} is the background signal when [target] = 0. In all modelling calculations herein, $I_{max} = 2.5 \times 10^5$ and $I_{bg} = 58$. These values were based on the approximate maximum binding signal and background signal that was measured in the flow cytometry assay.

Note that the model described herein is effectively a Langmuir model. However, in contrast to the normal Langmuir surface adsorption scenario where it is assumed that there is a large excess of potential adsorbates, the total quantity of surface sites in our system is comparable to the total quantity of potential adsorbates. Therefore, it was impractical to use the concentration of unbound target as an independent variable.

The ten possible sequence pairs when considering two mutations (SNP_1 and SNP_2) on two gene copies (target_1 and target_2) are as follows:

- $\text{target}_1 = SNP_1/SNP_2$ and $\text{target}_2 = SNP_1/SNP_2$ (homozygous double, SNP_1 and SNP_2)
- $\text{target}_1 = SNP_1/SNP_2$ and $\text{target}_2 = SNP_1/WT_2$

- $target_1 = SNP_1/SNP_2$ and $target_2 = WT_1/SNP_2$
- $target_1 = SNP_1/SNP_2$ and $target_2 = WT_1/WT_2$ (heterozygous *cis*)
- $target_1 = SNP_1/WT_2$ and $target_2 = SNP_1/WT_2$ (homozygous single, SNP_1)
- $target_1 = SNP_1/WT_2$ and $target_2 = WT_1/SNP_2$ (heterozygous *trans*)
- $target_1 = SNP_1/WT_2$ and $target_2 = WT_1/WT_2$
- $target_1 = WT_1/SNP_2$ and $target_2 = WT_1/SNP_2$ (homozygous single, SNP_2)
- $target_1 = WT_1/SNP_2$ and $target_2 = WT_1/WT_2$
- $target_1 = WT_1/WT_2$ and $target_2 = WT_1/WT_2$ (wild-type)

The equations for total binding affinity to the SNP_1/SNP_2, SNP_1/WT_2, WT_1/SNP_2 and WT_1/WT_2 targets can be described as

$$\text{Total } K_{\text{eq,SNP1/SNP2}} = K_{\text{eq,S1}} + K_{\text{eq,S2}} + K_{\text{eq,S1}} \times K_{\text{eq,S2}} \times c_{\text{eff}}$$

$$\text{Total } K_{\text{eq,SNP1/WT2}} = K_{\text{eq,S1}} + \text{MM}_2 \times K_{\text{eq,S2}} + K_{\text{eq,S1}} \times \text{MM}_2 \times K_{\text{eq,S2}} \times c_{\text{eff}}$$

$$\text{Total } K_{\text{eq,WT1/SNP2}} = \text{MM}_1 \times K_{\text{eq, S1}} + K_{\text{eq,S2}} + \text{MM}_1 \times K_{\text{eq,S1}} \times K_{\text{eq,S2}} \times c_{\text{eff}}$$

$$\text{Total } K_{\text{eq,WT1/WT2}} = \text{MM}_1 \times K_{\text{eq, S1}} + \text{MM}_2 \times K_{\text{eq,S2}} + \text{MM}_1 \times K_{\text{eq,S1}} \times \text{MM}_2 \times K_{\text{eq,S2}} \times c_{\text{eff}}$$

where MM_1 corresponds to a mismatch in S_1' and MM_2 corresponds to a mismatch in S_2'.

Synthesis of DNA-functionalized silica particles

The 5-µm amine-modified silica particles were suspended in dimethyl sulfoxide (DMSO) at a concentration of ~4.4×10^5 particles µl^{-1}. A 100 mg ml^{-1} stock of succinimidyl 4-(*N*-maleimidomethyl) cyclohexane-1-carboxylate was then prepared in DMSO and added at a final concentration of 10 mg ml^{-1} to a solution of 4×10^4 particles µl^{-1} suspended in DMSO. The reaction was then incubated at room temperature for 30 min to prepare maleimide-labelled silica beads. During the reaction of succinimidyl 4-(N-maleimidomethyl) cyclohexane-1-carboxylate with the amine-modified silica beads, 0.5 µl of 100 µM thiolated DNA (1 µM final), 0.5 µl of 10 mM tris(2-carboxyethyl)phosphine (TCEP; 100 µM final) and 49 µl of 1× phosphate buffered saline (PBS) at pH 6.8 were mixed and incubated at room temperature for 30 min to 1 h to reduce the thiolated DNA. The maleimide-labelled silica bead solution was then centrifuged on a tabletop mini-centrifuge at 2,000*g* for 1 min and the supernatant was removed and replaced with an equal volume of a 1 mg ml^{-1} solution of sulfo-*N*-hydroxysuccinimide (NHS) acetate dissolved in DMSO. The reaction was then incubated at room temperature for 30 min to prepare passivated, maleimide-labelled silica beads. Following the 30-min incubation, centrifugation and supernatant removal was performed four times. After each of the first three centrifugations, the beads were resuspended in a 2× volume of DMSO. Following the third resuspension, the beads were split into 50 µl aliquots and then centrifuged for a fourth time. After the fourth centrifugation, the supernatant was removed and replaced with the 50 µl solution of 1 µM reduced thiol-DNA and incubated at room temperature overnight or >8 h. For $n = 2$ beads, a pre-mixed solution containing 500 nM of each oligo was added, resulting in a total concentration of 1 µM of DNA. Following incubation, 50 µl of 1× saline sodium citrate (SSC), 0.1% Tween20 was added to each tube to help with centrifugation. Next, centrifugation and supernatant removal was performed four times. After each of the first three centrifugations, the beads were resuspended in 100 µl of 1× SSC, 0.1% Tween20. After the fourth centrifugation, the beads were resuspended in 1 ml of 1× SSC, 0.1% Tween20.

Determining the number of oligos per silica particle

The approximate concentration of the DNA-coated bead stocks was determined from four representative stock solutions using a haemacytometer. Then, two volumes containing ~1×10^5 or ~1.75×10^5 beads were taken from the four different DNA-coated bead stocks. Centrifugation at 2,000*g* with the tabletop mini-centrifuge and supernatant removal was then performed for each sample, followed by resuspension in 100 µl of 0.1 M KOH to dissolve the beads[45]. The beads were incubated in the KOH solution at room temperature for > 8 h. Bright-field microscopy images before and after KOH incubation were obtained using a Rebel bright-field microscope (Echo). The bead solution was then centrifuged again, and the supernatant was removed and added to a new tube. The centrifuged bead solution was resuspended in 1× SSC, 0.1% Tween20 and analysed using flow cytometry to confirm that the beads were etched/dissolved fully. The tube containing the removed supernatant was filtered using P2 gel filtration to remove KOH from the solution, then 20× Tris-EDTA (TE) buffer was added to the solution to give a final 1× concentration of TE buffer. Samples were then transferred to a 96-well plate. Oligreen was added to the solution at a final concentration of 1× and incubated for ~5 min at room temperature before the fluorescence was measured using a Biotek plate reader with Gen5 Microplate Reader and Imager software version 3.08. To generate a standard curve of fluorescence versus [DNA], 0, 5, 10, 20, 35 and 50 nM solutions of unreduced thiol DNA were prepared and incubated in a solution of 0.1 M KOH for >1 h. Following KOH incubation, the DNA solution was filtered using P2 gel filtration to remove KOH from the solution. A 20× TE buffer was then added to the solution to give a final 1× concentration of TE buffer. Samples were then transferred to a 96-well plate. Oligreen was added to the solution at a final concentration of 1× and incubated for ~5 min at room temperature before the fluorescence was measured using the Biotek plate reader. Using the standard curve, the concentration of DNA in the bead samples was determined and then divided by initial bead concentration to determine the number of oligos per silica bead.

Atto647N conjugation to target strands

Excess NHS-Atto647N (250 µg) was dissolved in 10 µl of fresh DMSO and then added to 10 nmol of amine-labelled target strands in 1× PBS with 0.1 M $NaHCO_3$. The reaction was left for >1 h at room temperature. After incubation, unreacted NHS-Atto647N and salts were removed by P2 or P4 gel filtration and purified using analytical-scale reverse-phase high-performance liquid chromatography (HPLC) with an Agilent AdvanceBio Oligonucleotide C18 column and OpenLAB CDS ChemStation edition software. The product was eluted in solvents A (0.1 M triethlyamine acetate (TEAA)) and B (acetonitrile (ACN)) with linear gradients of 10–35% solvent B over 25 min and 35–100% solvent B over 5 min at 0.5 ml min^{-1} flow rate. The molecular weight of the products was evaluated with an electron spray ionization (ESI) method using a Thermo Fisher Scientific Orbitrap system. The samples were prepared in a mixture of 70% Nanopure water and 30% methanol containing 10 µM EDTA, 0.0375% triethylamine and 0.75% 1,1,1,3,3,3-hexafluoro-2-propanol, and the spectra were recorded in negative charge mode eluted with the same solution[46]. The main peak of the obtained ESI-MS spectrum (m/z) was then deconvoluted to obtain the average molecular weight for the oligonucleotides. The concentration of the strands was determined by UV–vis using a Nanodrop instrument.

Flow cytometry assay and analysis to measure target binding

For all data except that shown in Fig. 6g,h and Extended Data Fig. 10, 1 nM Atto647N-labelled target was added to ~2.5×10^4 DNA-coated silica beads suspended in 1× SSC, 0.1% Tween20 and incubated at room temperature for 1 h. For the data in Fig. 6g,h, 10 nM Atto647N-labelled target was used. For the data in Extended Data Fig. 10, 1, 5, 10, 25 and 100 nM Atto647N-labelled target was used. Following the 1-h incubation, centrifugation at 2,000*g* with the tabletop mini-centrifuge and

supernatant removal was performed four times. After each of the four centrifugations, the beads were resuspended in 100 µl of 1× SSC, 0.1% Tween20. Fully washed beads were then injected into the flow cytometer for final data collection using CytExpert 2.3. After performing flow cytometry, FlowJo V10 was used to analyse the data. Singlet beads were isolated from the sample by gating first using forward scatter and side scatter area and second using forward and side scatter height. The MFI of the singlet beads from each sample was then calculated and plotted using GraphPad Prism 9. To account for higher non-specific binding due to the higher target concentrations used, MFIs for the data in Fig. 6g,h and Extended Data Fig. 10 were subtracted by the difference between the MFI of a non-complementary bead functionalized with the 4T oligo binding the same target at the same concentration and the background MFI (58) of the assay.

Fluorescence microscopy of Atto647N-labelled targets hybridized to beads

Fluorescence microscopy was also used to image targets hybridized to the beads and confirmed homogeneous binding across the bead surface (Supplementary Fig. 6). For this experiment, wells in a glass-bottom 96-well plate were soaked in ethanol for 5 min, rinsed with Nanopure water, and coated in 10% bovine serum albumin (BSA) for 15 min and rinsed again before imaging, then 0 or 1 nM Atto647N-labelled target was added to ~2.5×10^4 DNA-coated silica beads suspended in 1× SSC, 0.1% Tween20 and incubated at room temperature for >1 h. Following the incubation, centrifugation at 2,000*g* with the tabletop mini-centrifuge and supernatant removal was performed four times. After each of the four centrifugations, the beads were resuspended in 100 µl of 1× SSC, 0.1% Tween20. Fully washed beads were then added to the 96-well microscopy plate and imaged on the fluorescence microscope. Bright-field images were obtained and Atto647N images were acquired using a Cy5 cube. Images were obtained using NIS-Elements software and processed using Fiji ImageJ software.

Preparation of virus-extracted SARS-CoV-2 targets

Heat-inactivated original- and omicron-strain SARS-CoV-2 particles were provided by the NIH RADx-Radical Data Coordination Center (DCC) at the University of California San Diego and BEI Resources and by A. Carlin and the UCSD CALM and EXCITE laboratories, respectively. Viral RNA was extracted using a QIAamp Viral RNA mini kit from Qiagen following the accompanying protocol. RNA was aliquoted and stored at −80 °C until use. For both the original and omicron RNA samples, 1,000 ng of RNA was reverse-transcribed into cDNA using a High-Capacity cDNA Reverse Transcription kit from Thermo Fisher in a T100 Thermal Cycler (BioRad). The cDNA was then amplified using asymmetric PCR with an Atto647N-labelled forward primer and an unmodified reverse primer at 20 µM and 2 µM, respectively. PCR amplification was performed using the GeneAmp Fast PCR Master Mix kit from Thermo Fisher using the following 'three-step' PCR protocol:

1. 95 °C, 30 s
2. 95 °C, 3 s
3. 54 °C, 15 s
4. 72 °C, 10 s
5. Repeat steps 2–4 59 times (60 cycles total)
6. 65 °C, 1 min
7. Ramp to 97 °C at 0.2 °C s^{-1}
8. Hold at 4 °C

Viral products were purified using an analytical-scale reverse-phase HPLC with an Agilent AdvanceBio Oligonucleotide C18 column. The products were eluted in solvents A (0.1 M TEAA) and B (ACN) with linear gradients of 10–35% solvent B over 25 min and 35–100% solvent B over 5 min at a flow rate of 0.5 ml min^{-1} and then verified using native-polyacrylamide gel electrophoresis (PAGE). A 15% native PAGE gel was made with 1× TBE buffer and was run at 200 V for 1.5 h. Next, the gel was stained with 0.5 µg ml^{-1} ethidium bromide for 10 min and imaged using a fluorescent gel imager (Invitrogen). The concentration of the strands was determined by UV–vis using a Nanodrop instrument.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this Article.

Data availability

Data files containing all individual replicate data from main text figures and all individual replicate data from Extended Data figures are provided with this manuscript. All *P* values from the performed statistical analyses are provided in the corresponding figure captions. Source data are provided with this paper.

References

45. Bossert, D. et al. A hydrofluoric acid-free method to dissolve and quantify silica nanoparticles in aqueous and solid matrices. *Sci. Rep.* **9**, 7938 (2019).
46. Hail, M., Elliott, B. & Anderson, K. High throughput analysis of oligonucleotides using automated electrospray ionization mass spectrometry. *Am. Biotechnol. Lab.* **12**, 12–14 (2004).

Acknowledgements

Heat-inactivated SARS-CoV-2 and human corona (229E, OC43) virus samples (original/Washington strain) at known concentrations were provided by the NIH RADx-Radical Data Coordination Center (DCC) at the University of California San Diego and BEI Resources. The following reagent was obtained from UCSD: SARS-related coronavirus 2, isolate hCoV-19/USA/CA-SEARCH-59467/2021 (lineage BA.1; omicron variant), contributed by A. Carlin and the UCSC CALM and EXCITE laboratories. We acknowledge support from grant no. NSF MSN 2004126. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

B.R.D., R.M., S.N., H.O. and Y.D. planned and carried out the experiments. J.T.K. designed and derived the mathematical model. B.R.D., R.M. and K.S. conceived the original idea. B.R.D. analysed the data and took the lead in writing the manuscript. All authors provided critical feedback and helped shape the research, analysis and manuscript.

Competing interests

The authors declare no competing interests.

Additional information

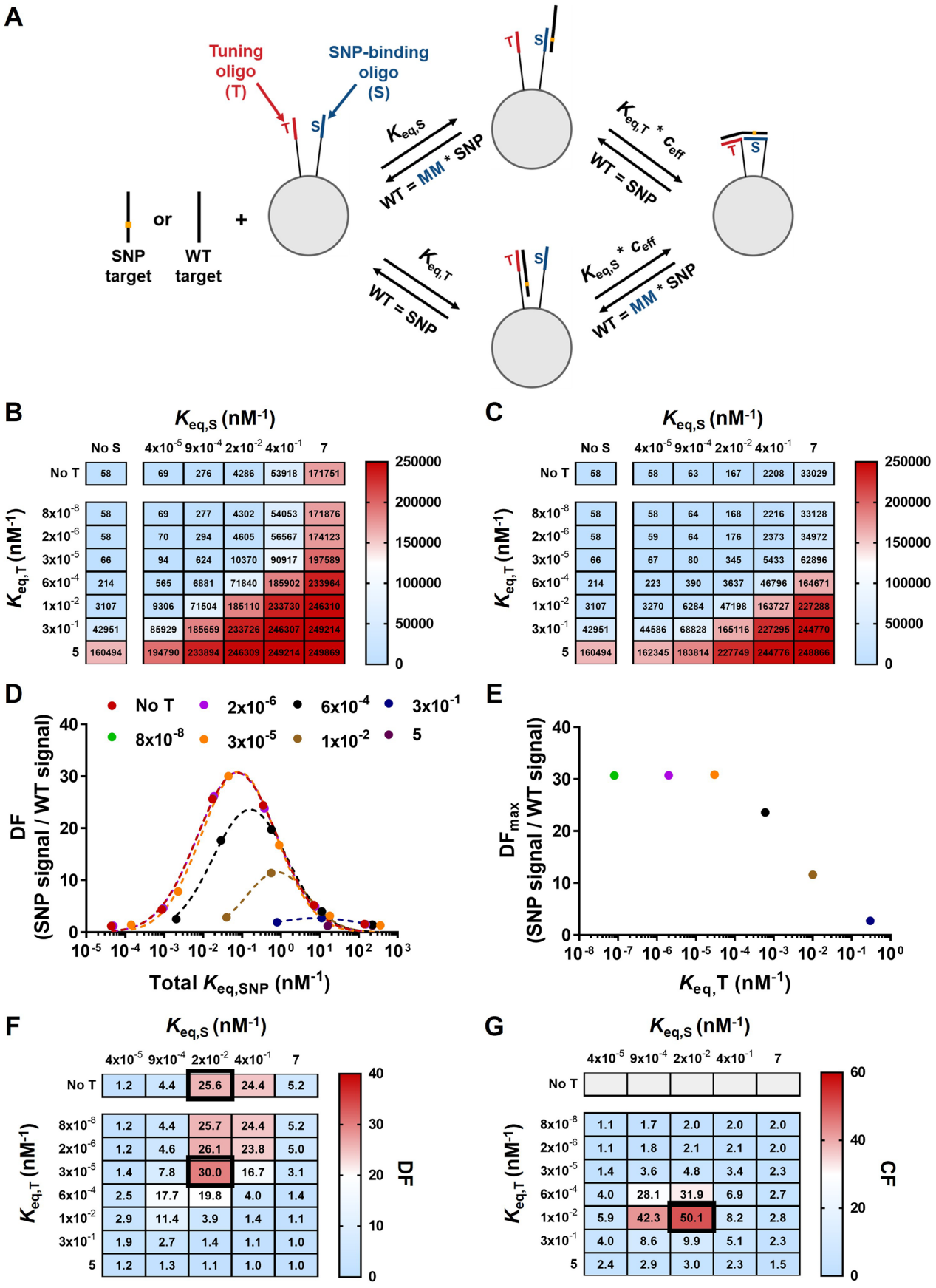
Extended data is available for this paper at https://doi.org/10.1038/s41557-023-01345-4.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41557-023-01345-4.

Correspondence and requests for materials should be addressed to Khalid Salaita.

Peer review information *Nature Chemistry* thanks Tine Curk, Jessica Rouge and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

https://doi.org/10.1038/s41557-023-01345-4

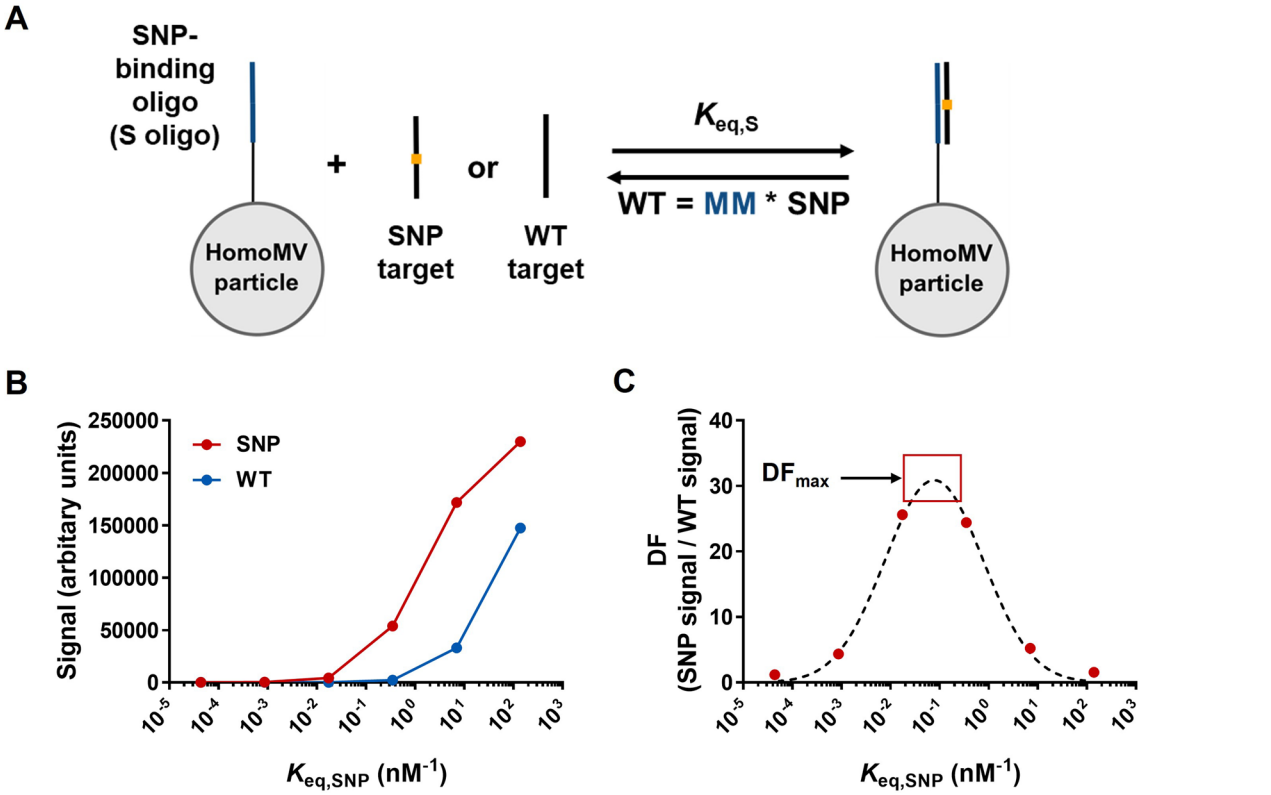


Extended Data Fig. 1 | See next page for caption.

https://doi.org/10.1038/s41557-023-01345-4

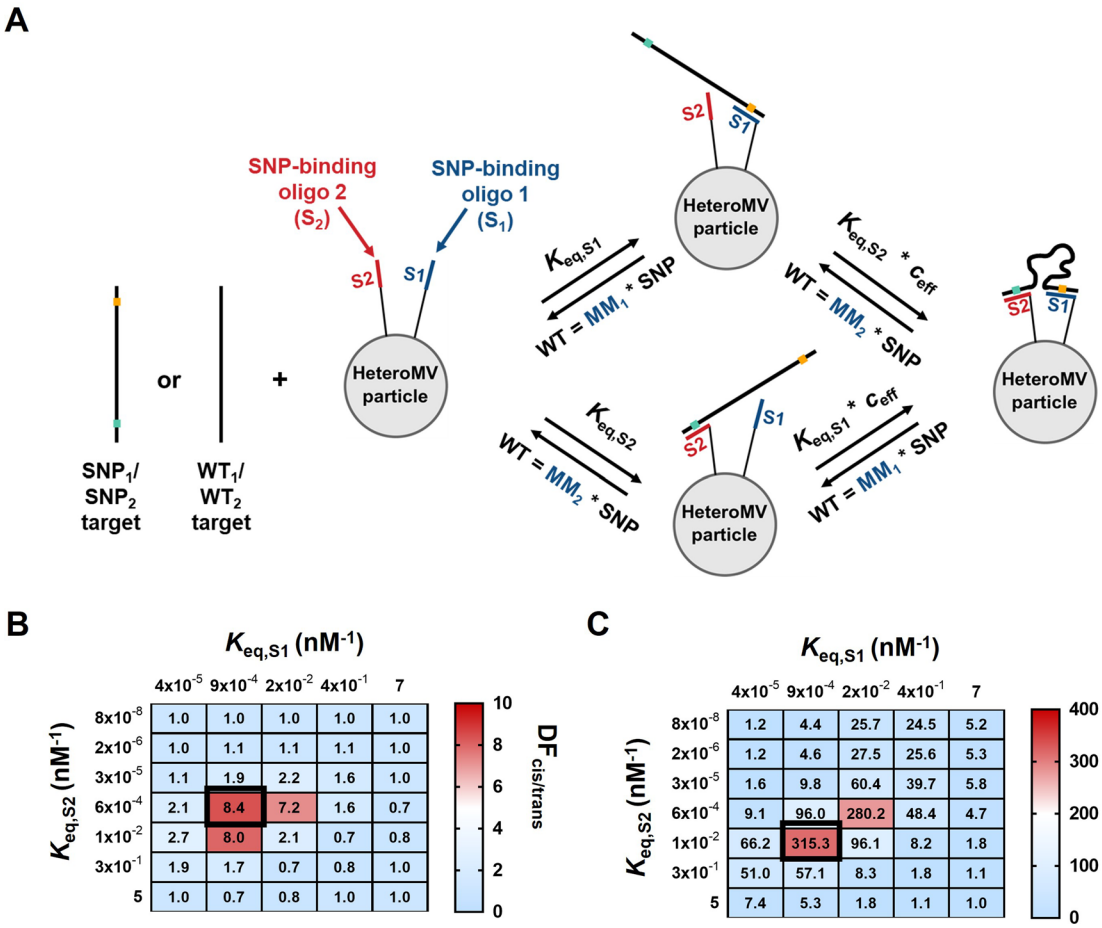
Extended Data Fig. 1 | Modelling the specificity of heteroMV particles for single mutant targets. a, Scheme showing the two-step reversible binding pathway of an $n = 2$ heteroMV particle binding either a SNP-containing target or a WT target and corresponding equations used to model the binding affinity to each target. **b, c**, Heatmap showing the predicted arbitrary signals when binding the SNP target (**b**) or WT target (**c**) as the monovalent binding affinities of the S and the T oligo are varied. **d,** Predicted discrimination factors for an $n = 2$ heteroMV particle as the affinity of T ($K_{eq,T\ only}$) is increased (different color dots) causing the total affinity for the SNP target ($K_{eq,S+T,SNP}$) to increase (x-axis) for each discrete value of $K_{eq,S\ only}$ chosen (same color dots). The curves were generated by fitting the predicted values to a gaussian distribution in GraphPad. **e,** The maximum DF value predicted from the curve in (**d**) for each discrete value of $K_{eq,T\ only}$. **f,** Heatmap showing the predicted discrimination factor when the monovalent binding affinities of the S and the T oligo are varied. Black boxes indicate the $n = 1$ and $n = 2$ combination with the highest discrimination factors. **g,** Heatmap showing the predicted cooperativity factor when the monovalent binding affinities of the S and the T oligo are varied. Black box indicates the $n = 2$ combination with the highest cooperativity factor.

https://doi.org/10.1038/s41557-023-01345-4



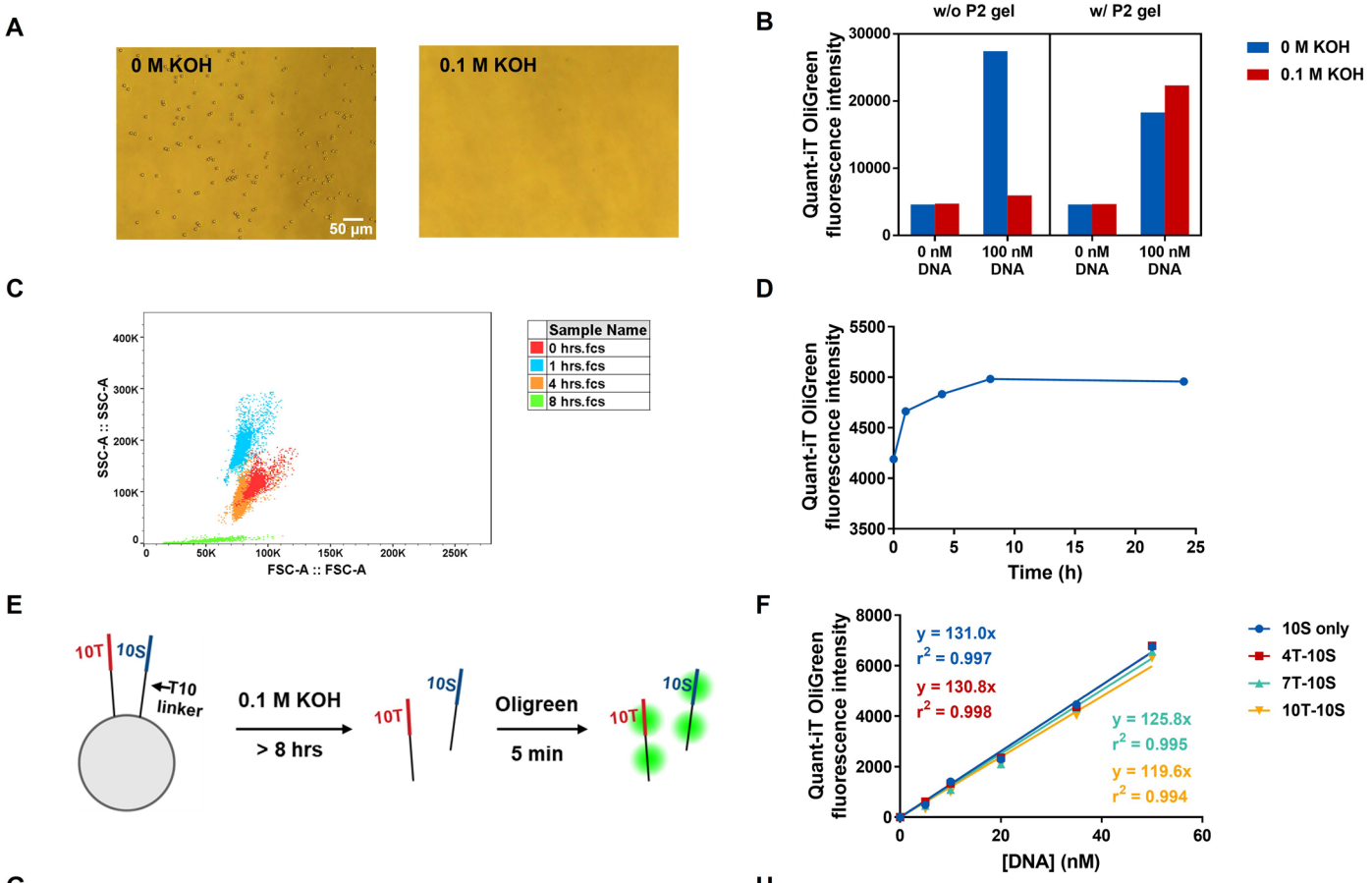
Extended Data Fig. 2 | Modelling the specificity of homoMV particles for single mutant targets. a, Scheme showing binding pathway of a homoMV particle binding either a SNP-containing target or a WT target and modification of binding affinity equation with MM factor. **b,** Predicted arbitrary signals when a homoMV particle with different affinities binds the SNP target or the WT target. **c,** Predicted discrimination factors for a homoMV particle with different affinities. Red dots correspond to discrimination factors for six values of K_{eq} chosen to mimic a series of oligos of length $x, x+1, \ldots, x+5$ nt. The black dashed curve was generated by fitting the predicted values to a gaussian distribution in GraphPad.

https://doi.org/10.1038/s41557-023-01345-4



Extended Data Fig. 3 | Modelling the specificity of heteroMV particles for double mutant targets. a, Scheme showing the two-step reversible binding pathway of an $n = 2$ heteroMV particle binding either a double mutant target or a double WT target and corresponding equations used to model the binding affinity to each target. **b,** Heatmap showing the predicted *cis*/*trans* discrimination factor when the monovalent binding affinities of the S_1 and the S_2 oligo are varied. Black box indicates the $n = 2$ combination with the highest predicted *cis*/*trans* discrimination factor. **c,** Heatmap showing the predicted double mutant discrimination factor when the monovalent binding affinities of the S_1 and the S_2 oligo are varied. Black box indicates the $n = 2$ combination with the highest predicted double mutant discrimination factor.

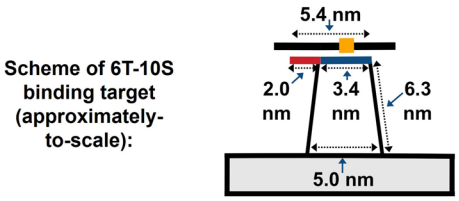
https://doi.org/10.1038/s41557-023-01345-4



G

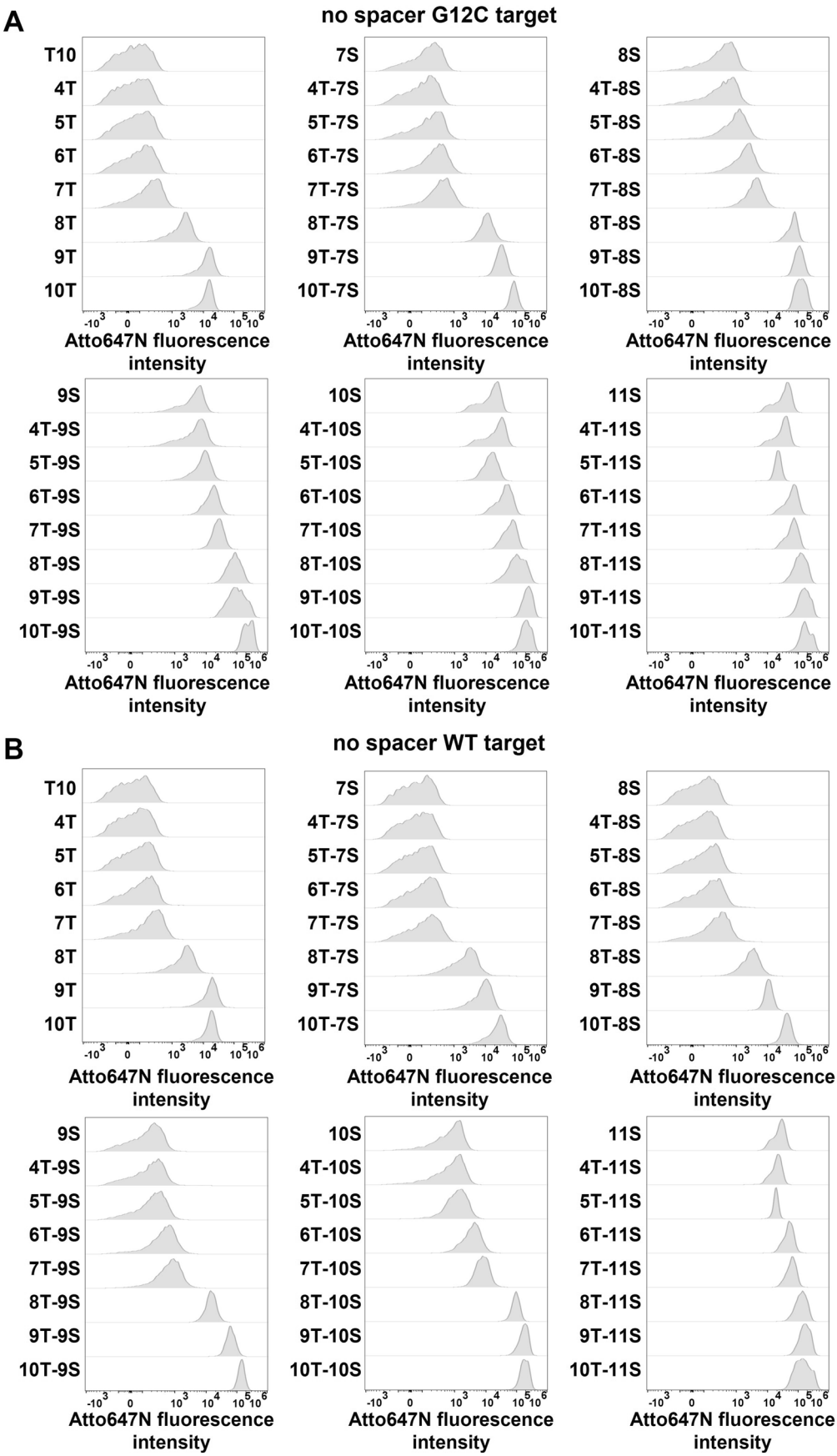
Bead name	# of beads	Total DNA concentration (nM)	DNA/um^2	distance between each oligo (nm)
10S only	9.4 x 10^4 1.6 x 10^5	4.46 8.84	3.6 x 10^4 4.1 x 10^4	5.2 4.9
4T-10S	9.0 x 10^4 1.6 x 10^5	5.55 8.88	4.7 x 10^4 4.3 x 10^4	4.6 4.8
7T-10S	1.2 x 10^5 2.0 x 10^5	6.09 8.98	4.0 x 10^4 3.4 x 10^4	5.0 5.5
10T-10S	8.7 x 10^4 1.5 x 10^5	3.76 9.89	3.3 x 10^4 5.0 x 10^4	5.5 4.5
Avg. ± SEM			4.1 ± 0.2 x 10^4	5.0 ± 0.1

H



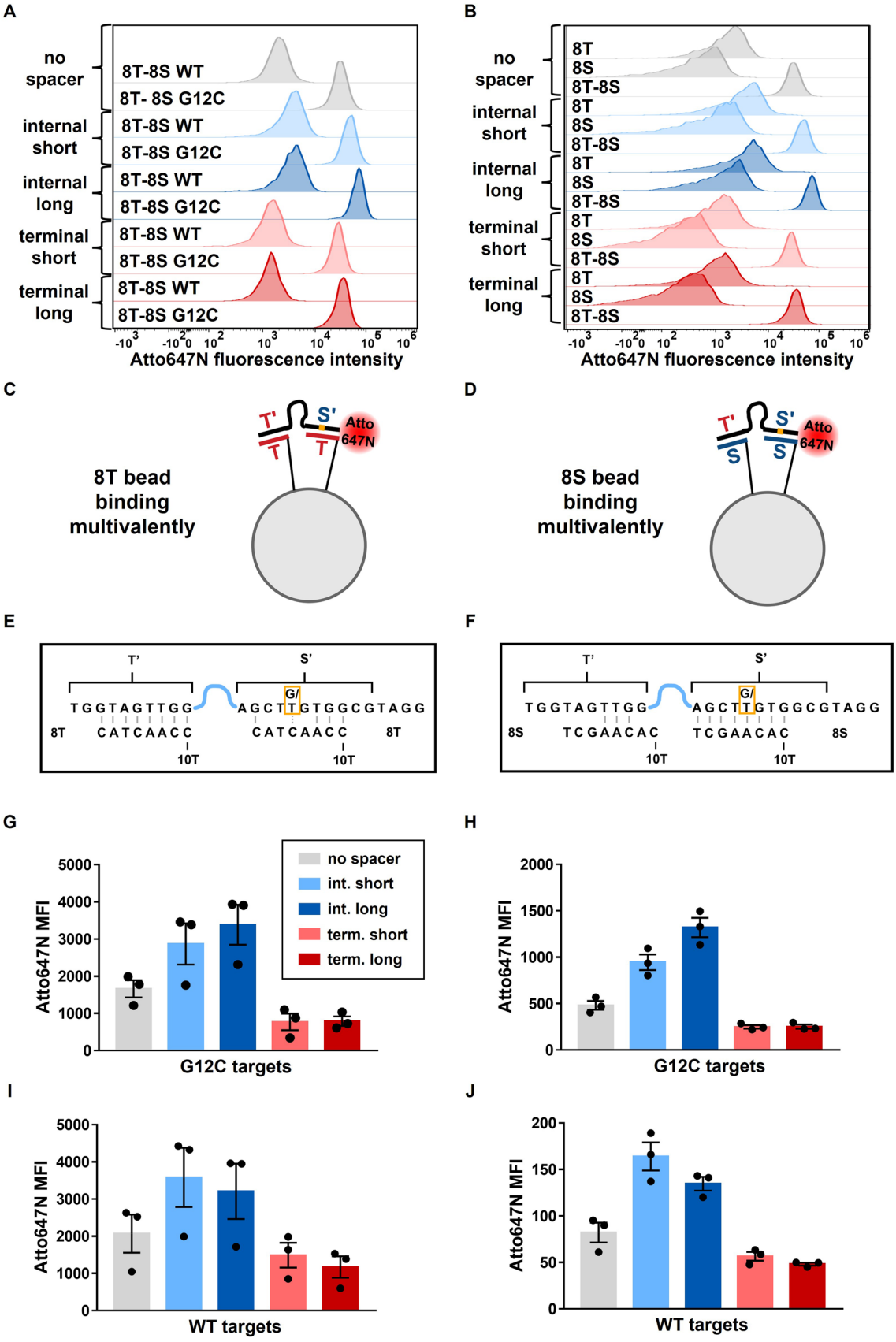
Extended Data Fig. 4 | Characterization of DNA-functionalized silica particles. a, Bright-field microscopy images of 5-µm silica beads incubated in 0 M or 0.1 M KOH overnight. Similar results were obtained from three independent experiments. **b,** Oligreen fluorescence intensity after incubation of 0 or 100 nM of a 20 nt oligo in 0 or 0.1 M KOH for ~6 hours. Following incubation, each sample was split into two tubes, and then Oligreen was added directly to the first tube and added following P2 gel filtration to the second tube. The plot shows that the presence of KOH in solution inhibits Oligreen fluorescence and that removing KOH using a P2 gel before adding Oligreen enables strong Oligreen fluorescence, though some DNA may be lost during filtration. **c,** Flow cytometry plot showing side scatter vs forward scatter area of DNA-coated 5-µm silica beads after incubation in 0.1 M KOH for 0, 1, 4, or 8 hours. The plot shows that over time, the bead size is reduced, and the bead structure is damaged following KOH incubation, suggesting that the DNA has been released from the bead surface. **d,** Oligreen fluorescence intensity following incubation of beads in 0.1 M KOH for 0, 1, 4, 8, or 24 hours, followed by P2 gel filtration. The plot indicates that all of the DNA has been released from the beads after ~8 hours. **e,** Scheme showing the finalized assay for quantifying the density of the DNA on the silica beads using the Oligreen reagent. **f,** Standard curves of Oligreen fluorescence intensity vs [DNA] from different concentrations of 4 different oligo mixtures. **g,** Table showing the # of beads, the measured [DNA], DNA/µm^2 on the bead surface, and the calculated mean distance ± standard error of the mean between each oligo on the bead surface from the $n = 8$ distinct samples. **h,** Approximately-to-scale illustration of a 6R-10S bead binding the no spacer G12C target based on the DNA density measurements and literature values for single stranded and double stranded DNA lengths.

https://doi.org/10.1038/s41557-023-01345-4



Extended Data Fig. 5 | Representative histograms for all bead combinations binding the no spacer targets. a,b, Representative histograms for all bead combinations binding the no spacer G12C target (**a**) and WT target (**b**) with one plot for each heatmap column from Fig. 2c and d.

https://doi.org/10.1038/s41557-023-01345-4

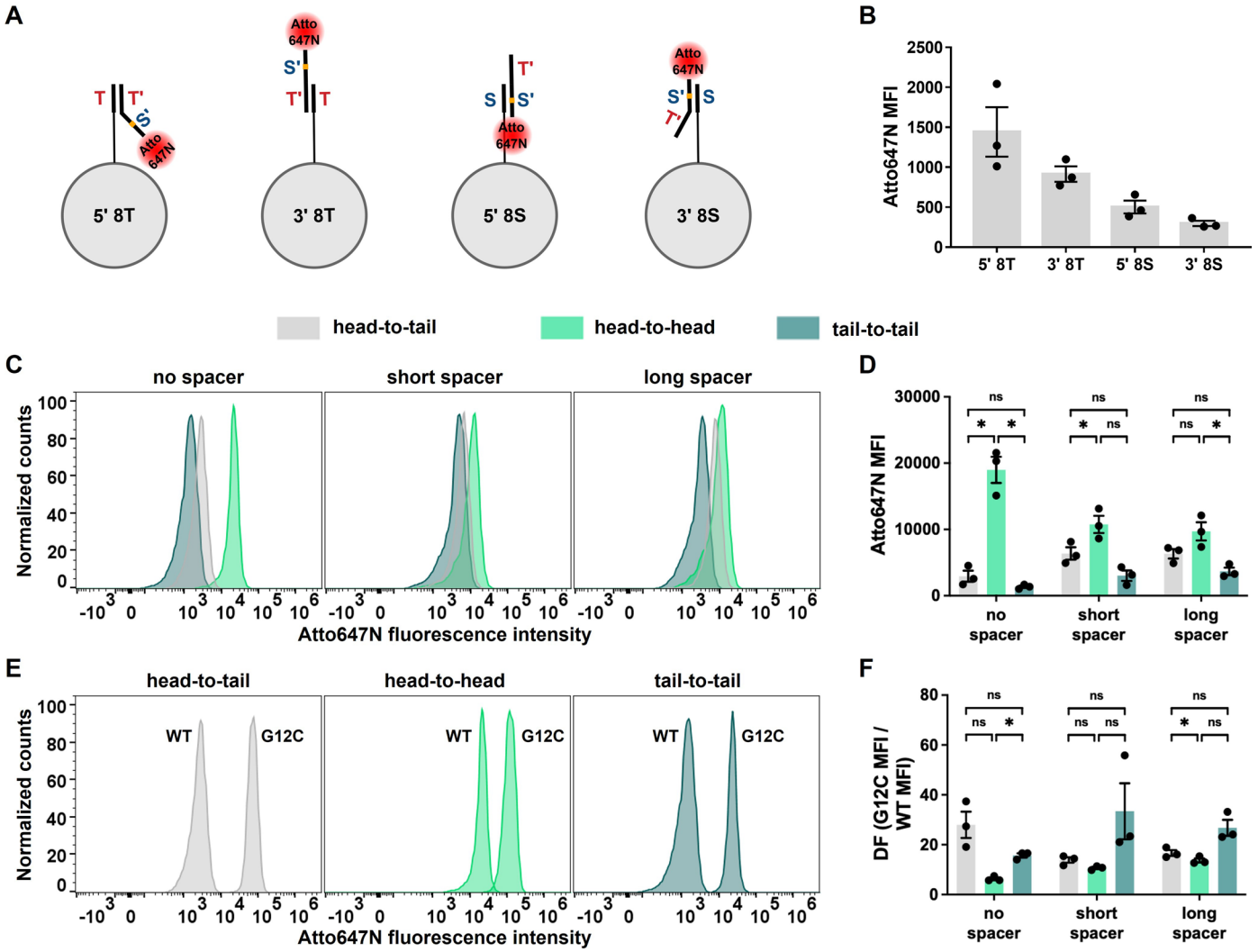


Extended Data Fig. 6 | See next page for caption.

https://doi.org/10.1038/s41557-023-01345-4

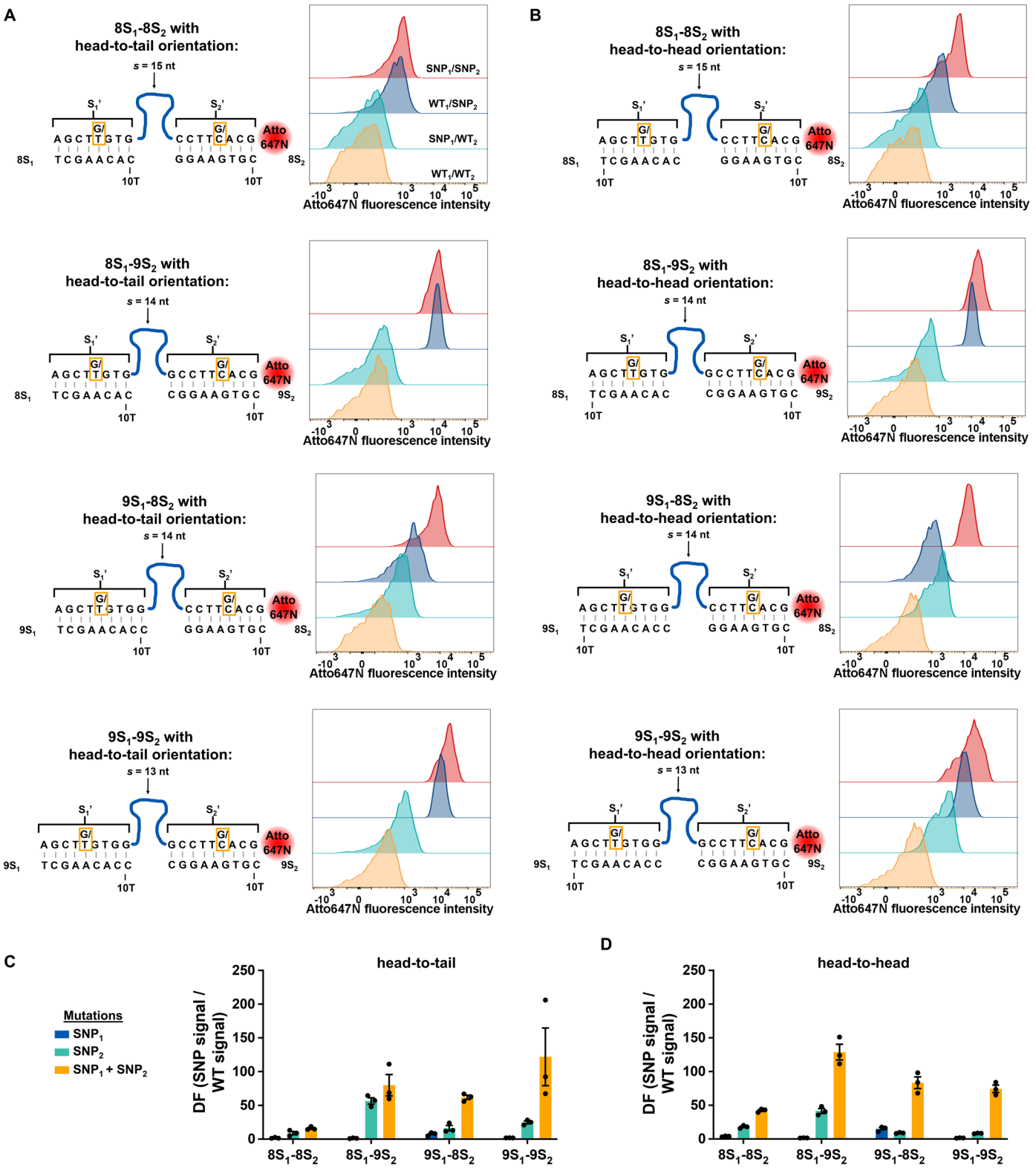
Extended Data Fig. 6 | Impact of spacer length and type on binding of 8T, 8S, and 8T-8S beads. a, Representative histograms for the 8T-8S beads binding the WT or G12C version of each of the different spacer length and spacer type targets. **b,** Representative histograms for the 8T, 8S, and 8T-8S beads binding the G12C version of each of the different spacer length and spacer type targets. Surprisingly, 8T and 8S only beads also showed increased binding to the internal spacer-containing targets, potentially due to weak binding between S′ and T as well as T′ and S. **c,d,** A simplified hypothetical illustration showing the 8T (**c**) and 8S (**d**) beads binding multivalently to an internal spacer-containing target. **e,f,**) Scheme showing the possible base pairs formed for the 8T (**e**) and 8S (**f**) beads binding multivalently to the target. **g–j,** Measured median fluorescence intensity values for the 8T (**g,i**) and 8S (**h,j**) beads binding the G12C (**g,h**) or WT (**i,j**) version of each of the different spacer length and spacer type targets. Error bars represent standard error of the mean from $n = 3$ distinct samples.

https://doi.org/10.1038/s41557-023-01345-4



Extended Data Fig. 7 | Impact of linker orientation on *n* = 1 bead binding and representative histograms for *n* = 2 beads binding WT targets. a, Scheme illustrating the possible binding interaction of the 5′ 8T, 3′ 8T, 5′ 8S, and 3′ 8S beads binding the no spacer G12C target monovalently. **b,** Measured median fluorescence intensity values for the 5′ 8T, 3′ 8T, 5′ 8S, and 3′ 8S beads binding the no spacer G12C target. As expected, the oligo's anchoring terminus did not have a notable effect on *n* = 1 beads binding the G12C no spacer target. **c,d,** Representative histograms (**c**) and measured median fluorescence intensity values (**d**) for 8T-8S beads with head-to-tail, head-to-head, or tail-to-tail orientation binding the WT target with no spacer, short spacer, or long spacer (*P* values for no spacer: head-to-tail vs. head-to-head = 0.0126, head-to-tail vs. tail-to-tail = 0.2359, head-to-head vs. tail-to-tail = 0.0185; *P* values for short spacer: head-to-tail vs. head-to-head = 0.0131, head-to-tail vs. tail-to-tail = 0.2934, head-to-head vs. tail-to-tail = 0.0997; *P* values for long spacer: head-to-tail vs. head-to-head = 0.1217, head-to-tail vs. tail-to-tail = 0.1463, head-to-head vs. tail-to-tail = 0.0469). **e,** Representative histograms for 8T-8S beads with head-to-tail, head-to-head, or tail-to-tail orientation binding the G12C and WT no spacer targets. **f,** Measured discrimination factors for 8T-8S beads with head-to-tail, head-to-head, or tail-to-tail orientation binding the no spacer, short spacer, or long spacer targets (*P* values for no spacer: head-to-tail vs. head-to-head = 0.082, head-to-tail vs. tail-to-tail = 0.2045, head-to-head vs. tail-to-tail = 0.012; *P* values for short spacer: head-to-tail vs. head-to-head = 0.0715, head-to-tail vs. tail-to-tail = 0.4181, head-to-head vs. tail-to-tail = 0.3223; *P* values for long spacer: head-to-tail vs. head-to-head = 0.0441, head-to-tail vs. tail-to-tail = 0.236, head-to-head vs. tail-to-tail = 0.1265). Error bars represent standard error of the mean from *n* = 3 distinct samples. Values were compared using paired one-way ANOVA with multiple comparisons follow-up tests ($^{ns}P > 0.05$, $^{*}P < 0.05$).

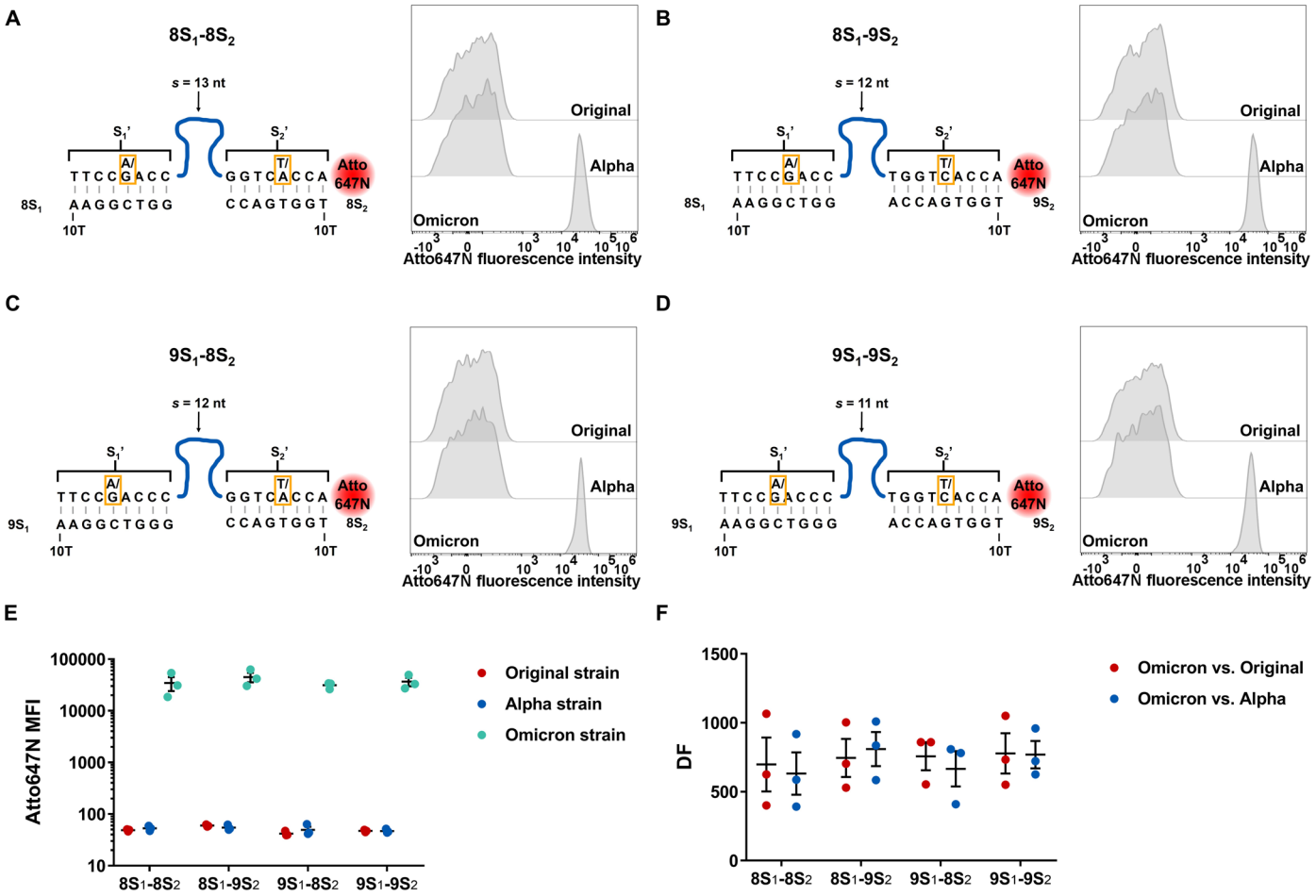
https://doi.org/10.1038/s41557-023-01345-4



Extended Data Fig. 8 | Representative histograms and results for all bead combinations binding the SNP_1/SNP_2, WT_1/SNP_2, SNP_1/WT_2, and WT_1/WT_2 targets. a,b, Scheme showing the sequences, anchor location, and spacer length for each bead combination with the head-to-tail orientation (**a**) or head-to-head orientation (**b**) and corresponding representative histograms for each bead combination binding the SNP_1/SNP_2, WT_1/SNP_2, SNP_1/WT_2, and WT_1/WT_2 targets. **c,d,** Measured discrimination factors for SNP_1, SNP_2, or $SNP_1 + SNP_2$ for each bead combination with the head-to-tail orientation (**c**) or head-to-head orientation (**d**). Error bars represent standard error of the mean from $n = 3$ distinct samples.

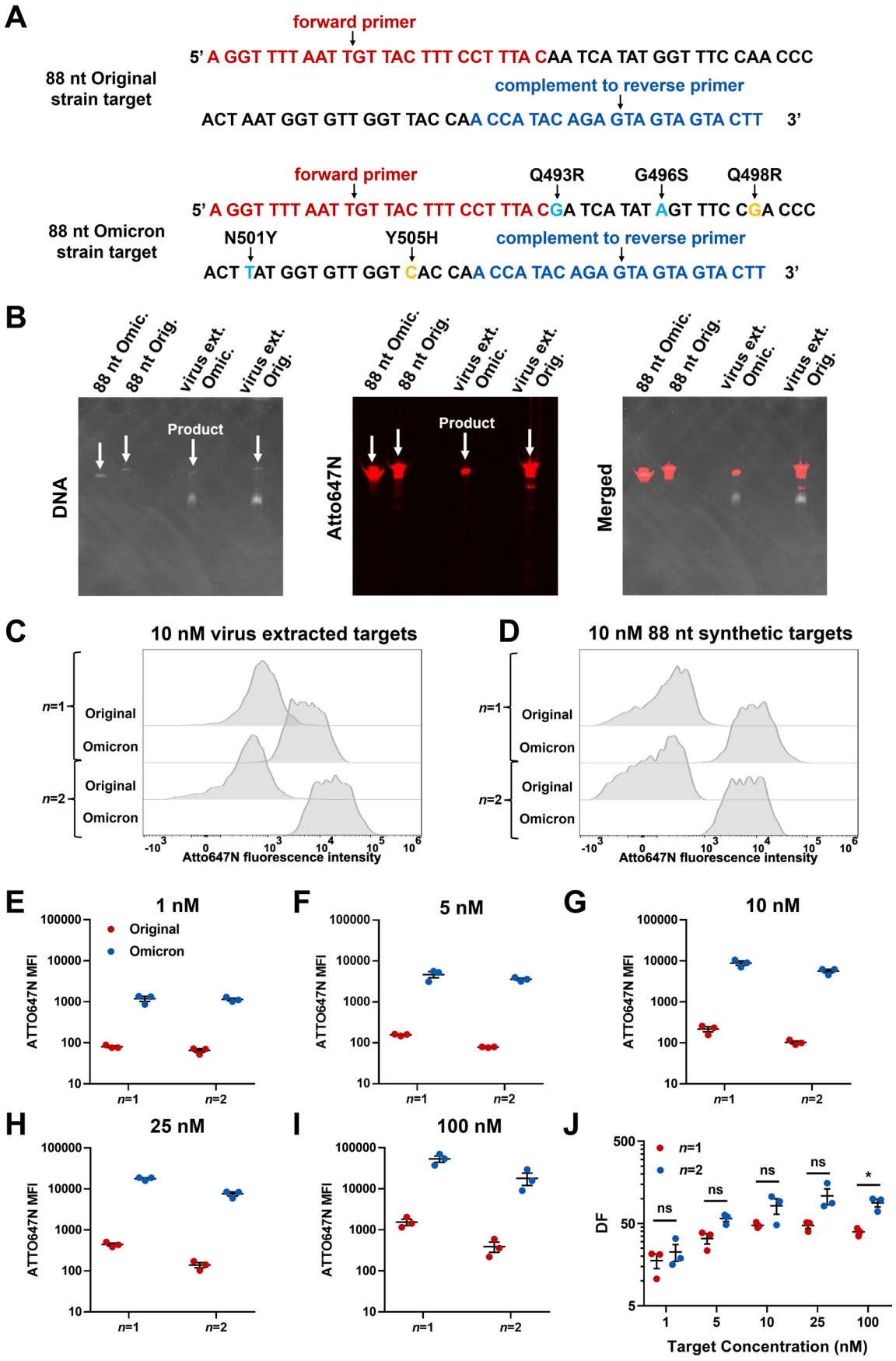
Interestingly, both beads containing the $8S_2$ oligo had higher $DF_{cis/trans}$ values when binding in the head-to-head orientation (Fig. 5g,h). Alternatively, beads containing the $9S_2$ oligo bound the SNP_1/SNP_2 and WT_1/SNP_2 targets similarly, resulting in poor specificity for SNP_1, and had similar $DF_{cis/trans}$ values in both orientations. This suggests that the $9S_2$ oligo's affinity for the target is too high resulting in a total binding affinity that is too strong to be appreciably impacted by a mismatch in the S_1' region. These results offer further evidence that the head-to-head orientation can yield higher binding, particularly when the two immobilized oligos are binding cooperatively.

https://doi.org/10.1038/s41557-023-01345-4



Extended Data Fig. 9 | Representative histograms and results for all bead combinations binding the model SARS-CoV-2 targets. a–d, Scheme showing the sequences, anchor location, and spacer length, as well as corresponding representative histograms for the $8S_1$-$8S_2$ (**a**), $8S_1$-$9S_2$ (**b**), $9S_1$-$8S_2$ (**c**), and $9S_1$-$9S_2$ (**d**) beads binding the original, alpha, and omicron strain targets. **e,** Measured median fluorescence intensity values for each bead combination binding the three targets. **f,** Measured discrimination factors for the omicron target versus the original or alpha targets for each bead combination. Error bars represent standard error of the mean from $n = 3$ distinct samples.

https://doi.org/10.1038/s41557-023-01345-4



Extended Data Fig. 10 | See next page for caption.

https://doi.org/10.1038/s41557-023-01345-4

Extended Data Fig. 10 | Design, characterization, and binding measurements for the 88 nt synthetic and SARS-CoV-2 virus extracted targets. a, Scheme showing the 88 nt original and omicron target sequences, as well as the mutations unique to the omicron target and the primer locations. The $9S_1$-$9S_2$ $n = 2$ bead is complementary to the Q498R and Y505H mutations (yellow) while the $n = 1$ bead is complementary to Q498R, N501Y, and Y505H. Therefore, the $n = 2$ bead binds the original target with two mismatches, whereas the $n = 1$ bead binds with three mutations. Primer sequences were designed to not overlap with any mutations, including the Q493R and G496S mutations. **b,** PAGE gel comparing synthetic 88 nt original and omicron targets to virus extracted original and omicron targets. The left image (DNA) shows ethidium bromide fluorescence, middle image (Atto647N) shows Atto647N fluorescence, and the right image (Merged) shows the two channels merged. **c,d,** Representative flow cytometry histograms for the $n = 1$ and $9S_1$-$9S_2$ $n = 2$ beads binding 10 nM virus extracted original and omicron targets (**c**) and 10 nM synthetic 88 nt original and omicron targets (**d**). Moreover, to assess the impact of target concentration on heteroMV binding specificity, synthetic versions of the same targets were tested at concentrations ranging from 1 to 100 nM. **e–i,** Measured median fluorescence intensity values for the $n = 1$ and $9S_1$-$9S_2$ $n = 2$ beads binding 1 nM (**e**), 5 nM (**f**), 10 nM (**g**), 25 nM (**h**), and 100 nM (**i**) of the synthetic 88 nt original and omicron targets. **j,** Measured discrimination factors corresponding to the data shown in **e**-**i** (*P* vales: 1 nM = 0.3583, 5 nM = 0.0907, 10 nM = 0.1717, 25 nM = 0.1071, 100 nM = 0.0224). $n = 2$ beads yielded increased average DF values at each concentration tested, showing that heteroMV beads can be used to improve specificity in a range of target concentrations. Error bars represent standard error of the mean from $n = 3$ distinct samples. Values were compared using two-sided paired student t tests ($^{ns}P > 0.05$, $^{*}P < 0.05$).

nature portfolio

Corresponding author(s): Khalid Salaita

Last updated by author(s): 6/22/2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	All flow cytometry data was collected using CytExpert 2.3 software from Beckman Coulter. All HPLC data was collected using OpenLAB CDS ChemStation edition from Agilent Technologies. All plate reader fluorescence measurements were collected using Gen5 Microplate Reader and Imager Software version 3.08 from BioTek. All microscopy images were collected using NIS-Elements version 4.50 from Nikon.
Data analysis	All flow cytometry data was analyzed using FlowJo V10. GraphPad Prism 9 was used to make all plots and to perform statistical analysis. Microsoft Excel 2016 was used for mathematical modeling calculations. All oligonucleotide secondary structure analysis was performed on NUPACK.org (free version). All microscopy images were analyzed using Fiji ImageJ 1.51s.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Source data files containing all individual replicate data from main text, all individual replicate data from supporting information, and all p values from statistical analyses performed are provided with this manuscript.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	All measurements were performed in at least triplicates from distinct samples. This sample size was chosen because it was considered sufficient to demonstrate the robustness and reproducibility of these measurements.
Data exclusions	No data was excluded from the described results.
Replication	All experimental findings were based on at least three distinct measurements. Relative trends between groups for each set of experiments were highly reproducible. Some variability in raw fluorescence values was observed from different DNA-functionalized bead synthesis batches and thus all compared data was produced from DNA-functionalized beads from the same batch, with the exception of the n=1 and n=2 DNA-functionalized beads compared in Figure 6C-6E. In this case, one side-by-side comparison (data not shown) was completed that was consistent with the data in Figure 6.
Randomization	All samples that were compared were tested side-by-side in the same experiment and thus no group allocation or randomization was necessary to investigate the parameters or perform the techniques described in this study.
Blinding	No group allocation was performed, therefore no blinding to group assignments was required.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Amine-capped 5 micrometer diameter silica beads were functionalized with thiolated oligonucleotides using SMCC hetero-bifunctional crosslinker. DNA-functionalized beads were then incubated with oligonucleotide targets, washed by a series of centrifugations and resuspensions, and loaded into the flow cytometer. See Methods section for further details.
Instrument	All flow cytometry data was collected using a CytoFLEX V0-B3-R1 flow cytometer equipped with a 488 nm and 638 nm laser from Beckman coulter
Software	All flow cytometry data was collected using CytExpert 2.3 software from Beckman Coulter and analyzed using FlowJo V10.
Cell population abundance	Median fluorescence intensities were calculated from approximately 5000 singlet DNA-functionalized beads, which comprised ~90% of the events analyzed per run on the flow cytometer. Singlet purity was determined based on forward and side scatter gating.
Gating strategy	For each sample, all events are first plotted on a side scatter area vs forward scatter area plot and a “singlets area” gate is drawn to approximately include just the singlet bead population. The cells included in the “singlets area” gate are then plotted on a side scatter height vs forward scatter height plot and a second gate, the “singlets height” gate, is drawn to more accurately include just the singlet bead population. See example plots in Figure S10.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.