

Efficient simulation of arbitrary multi-component first-order binding kinetics for improved assay design and molecular assembly

Kyle Briggs, Mohamed Yassine Bouhamidi, Liqun He, and Vincent Tabard-Cossa*

Department of Physics, University of Ottawa, Ottawa, Ontario, Canada

Traditional ELISA, long the workhorse for specific target protein detection using microplate wells, is nearing its fundamental limit of sensitivity. New opportunities in healthcare call for *in vitro* diagnostic tests with ultra-high sensitivity. Magnetic bead-based ELISA formats have been developed that can reach unprecedented sensitivities order of magnitude better than are allowed for by the rate constants for a single ligand-receptor interaction. However, these ultra-high sensitivity assays are highly vulnerable to a host of confounding factors, including nonspecific binding from background molecules and loss of low-abundance target to tube walls and during wash steps. Moreover, optimization of workflow is often time-consuming and expensive. In this work, we present a simulation tool that allows users to graphically define arbitrary binding assays, including fully reversible first-order binding kinetics, timed addition of extra components, and timed wash steps. The tool is freely available as a user-friendly webapp. The framework is lightweight and fast, allowing for inexpensive simulation and visualization of arbitrarily complex assay schemes, including but not limited to digital immunoassays, DNA hybridization, and enzyme kinetics, for validation and optimization of assay designs without requiring any programming knowledge from the user. We demonstrate some of these capabilities and provide practical guidance on assay simulation design.

The ability to sensitively detect specific biomarker in a clinical sample containing a mixture of off-target components is a cornerstone of diagnostic medicine. The most common method by which this is achieved for proteins is the enzyme-linked immunosorbent assay (ELISA),¹ in which target proteins are captured either by direct adsorption to the surface of a plate or through pre-coated “capture” antibodies, and subsequently detected by a labelled secondary “detector” antibody and optically read via colorimetric, fluorescence or chemiluminescence detection strategies.²⁻⁴ This technique is ubiquitous in the life sciences and medicine to detect and quantify a specific protein in a complex mixture and is driving everything from pregnancy tests to cancer detection. However, as our understanding of the human proteome advances, there is a growing need for detection of target proteins in the femtomolar concentration range, and standard ELISA is limited to target concentrations within a few logs of the dissociation constant (K_d) for the receptor-ligand interaction on which it is based, which practically limits it in most cases to the picomolar range.

Recently, impressive work has been done using so-called digital ELISA schemes, in which just a few copies of a target protein can be detected and counted directly.⁵⁻⁷ Digital counting methods overcome measurement uncertainty associated with integration of an analog optical signal and is in principle only limited by Poisson counting noise, making femtomolar and in some cases even attomolar concentrations detectable.^{8,9}

One of the best-known example of this approach is the SiMoA technology pioneered by D. Walt and now commercialized by Quanterix.^{5-7,10} In their scheme, paramagnetic beads are coated with capture antibodies specific to the target protein and are mixed with a clinical sample. Because each bead has on the

order of $N \sim 10^5$ antibodies, the effective K_d for a bead is N times smaller than the K_d for any individual antibody alone, permitting efficient capture in 3D of very low concentrations targets. Coupled with digital detection of the beads with or without a target protein bound, the SiMoA technology has improved the sensitivity of standard ELISA by $>1000\times$.¹¹

While this approach has made possible the quantification of low abundance biomarkers from complex biofluids,¹² the development of an assay for a particular target remains an extremely laborious and expensive task, involving weeks of optimization varying the assay steps, components concentrations, incubation times, number of washes, etc., to maximize the assay performance. To assist in these tedious experimental tasks, there is a need for simulation tools to shift the burden of optimization away from the expensive and time-consuming empirical framework.

Few studies have attempted to address this knowledge gap. Chang *et al*¹¹ presents a model that involves a multi-step, multi-component framework for simulation of their assay workflows which performs well in the low-concentration regime of target as compared to capture antibody (i.e. 1 or 0 targets per bead).¹¹ However, their model assumes irreversible binding between additions of new components to the mixture, which is overly optimistic about the final signal generated. More recently, a Python library that allows definition of arbitrary first-order coupled kinetics was introduced which has similar capabilities similar to the tool presented in this work, but relies on the user being able to program their own assays in Python.¹³ Other computational resources with slightly different specialization can be readily found,¹⁴ but applying these tools can be a laborious task.

In this work, we present a user-friendly tool that allows definition of arbitrarily complex multi-component systems of first-order binding interactions, which allows for timed addition of components part way through the workflow, and timed wash steps. The effects of nonspecific binding of background molecules and the influence of the timing and duration of wash steps can easily be simulated to inform assay design before undertaking complex experimental work. We demonstrate its utility through simulation of a full bead-based digital sandwich immunoassay, complete with nonspecific binding and wash steps, and discuss best practice for digital immunoassay design. We also model a binding reaction involving assembly of small DNA nanostructures used in our previous work^{15,16} with the goal of aiding in explaining previously ambiguous experimental results.

MATERIALS AND METHODS

Experimental. Solid-state nanopore data is reused from our previous work for the purposes of comparison to this tool. Briefly, nanopore fabrication is performed using the controlled breakdown method^{17–20} on SiN_x membranes and used to sense DNA nanostructures, the design of which has been previously published.²¹ Experiments are performed in 3.2 M LiCl pH 8, at 100 mV using a 12 nm pore, and a low-pass Bessel filtered at 200 kHz for analysis.

Theoretical. A simple reversible first-order receptor-ligand binding or DNA hybridization reaction between arbitrary components A and B is governed by the system of equations:

$$\frac{d[AB]}{dt} = k_{on}[A][B] - k_{off}[AB], \quad (1)$$

and

$$\frac{d[A]}{dt} = \frac{d[B]}{dt} = -k_{on}[A][B] + k_{off}[AB], \quad (2)$$

where $[\circ]$ denotes concentration and k_{on} and k_{off} are the association and dissociation constants which together define the dissociation constant $K_d = \frac{k_{off}}{k_{on}}$. This can be readily extended to a multi-component system with a vector of component concentrations \vec{c} using Einstein summation notation as:

$$\frac{dc^i}{dt} = \beta_j^{ik} c^j c_k + \alpha_j^i c^j, \quad (3)$$

where β is a rank-3 tensor of on-rates in such a way that $\beta_{ijk} = \beta_{ikj}$ for $i \neq j$ and $i \neq k$ is the rate at which c_i is produced by binding between c_j and c_k and β_{iik} for $i \neq k$ is the rate at which c_i is depleted as it forms various complexes with c_k . α is a matrix of off-rates such that α_{ij} for $i \neq j$ is the rate at which c_j breaks into c_i and another component, and α_{ii} is the total rate at which c_i is enriched by the breakup of all other components. To represent a physically valid system of first-order binding kinetics, α and β must satisfy conservation of mass requirements, which can be expressed as:

$$\alpha_{ii} = - \sum_{j \neq i} \alpha_{ji}, \quad (4)$$

and

$$\beta_{iik} = - \sum_{j \neq i} \beta_{jik}. \quad (5)$$

Prior to simulation, all variables are internally normalized in order to ensure numerical stability. We define two calculated normalization constants specific to a given system of interactions:

$$\bar{\beta} = \sqrt{\sum_{i,j,k} \beta_{ijk}^2} \quad (6)$$

is the on-rate normalization constant, while:

$$\bar{\alpha} = \sqrt{\sum_{i,j} \alpha_{ij}^2} \quad (7)$$

is the off-rate normalization constant. One could equivalently use the spectral norm of α to normalize, but we use the Frobenius norm here since it is easier to generalize to higher order tensors. From these, we define the population normalization, which can be thought of as a generalized analog of K_d for the system, as:

$$\kappa = \frac{\bar{\alpha}}{\bar{\beta}}. \quad (8)$$

Defining $\vec{\rho} = \kappa^{-1} \vec{c}$ and $\tau = \bar{\alpha} t$, $\mathbf{B} = \bar{\beta}^{-1} \beta$ and $\mathbf{A} = \bar{\alpha}^{-1} \alpha$, the normalized system of equations becomes:

$$\frac{d\rho^i}{d\tau} = \mathbf{B}_j^{ik} \rho^j \rho_k + \mathbf{A}_j^i \rho^j. \quad (9)$$

Note that all variables are rescaled back to the same units in which they are configured by the user prior to actual output.

While most non-trivial reactions do not have analytical solutions, this system can readily be numerically solved using an explicit Runge-Kutta integration scheme.²² The tool presented here uses standard RK4 to propagate Equation 9 through time, with a timestep of $\Delta\tau$ that is calculated at runtime to ensure numerical stability, under the assumption that the system will be farthest from equilibrium when the simulation starts. The timestep calculation will fail to ensure stability if the simulation is started close to an equilibrium configuration, which should never be the case for any practical model. The model outputs a snapshot of the current concentration vector 10 times per τ . Since τ is by construction shorter than the shortest dynamic timescale in the system this ensures that no interesting features are lost to sampling issues.

The model makes several key assumptions. We are considering only the evolution of concentration over time, which means that we are assuming that the probability of two species encountering one another is proportional to the ratio of their concentrations, or equivalently that all the components are perfectly mixed. We do not explicitly model diffusion, nor are steric interactions considered. As such it is important that this tool only be applied to simulation of assays that occur in the bulk phase (i.e., in 3D). If antibodies are immobilized on a 2D surface, for example, it is not expected that this model will give correct results.

For the purposes of numerical simulation, we define a system at equilibrium to be one that satisfies:

$$\left| \frac{d\rho}{d\tau} \right| \frac{\Delta\tau}{|\rho|} < 10^{-10}$$

The full user interface and a guide to using the tool for practical assay system and workflow definition is given in Supplementary Section S1. The webapp is freely available at immunoassay.herokuapp.com.

RESULTS AND DISCUSSION

Modelling First-Order Kinetics Binding Systems. We first begin with a demonstration of a reaction which can be solved analytically, namely, a simple ligand-receptor system defined by the reaction $A + B \leftrightarrow AB$. The results from the simulation are shown in Supplementary Section S2 to be in perfect agreement with theoretical predictions.

Idealized Digital Immunoassay. Digital immunoassay design is becoming an increasingly important challenge. As the limits of detection are pushed further down toward single molecule copies, the details of the sample preparation steps leading up to the final detection step become critical in enabling that level of sensitivity. To better understand the impact of upstream sample preparation and biochemical reactions leading up to the final sample read, we simulate a variety of model systems, including a full assay workflow used in the SiMoA technology to better understand the influence of the timing and duration of each step of a typical digital immunoassay.

The simulation setup for this model consists of four primary components: capture antibodies bound to magnetic beads (A), a target protein (T), a detector antibody (D), and a labelling molecule (L), such as an enzyme⁶ or a DNA strand.¹⁵ Assuming no cross-reactivity between species and no nonspecific background molecules, these four components can eventually bind together into an ATDL complex, forming every permutation of subcomplexes along the way. Our idealized assay consists of the 10 possible species representing the set of subcomponents: 4 base components (A, T, D, and L), 3 2-component subcomplexes (AT, TD, and DL), 2 3-component sub-complexes (ATD and TDL), and the full ATDL complex. These subcomponents can form or break up at any point. We will assume for simplicity that binding of a subsection of the complex does not alter the binding constants for subsequent

assembly steps, though the simulation framework can accommodate changing binding constants in response to partial complexation. This system is shown as a diagram in Figure 1a and 1b.

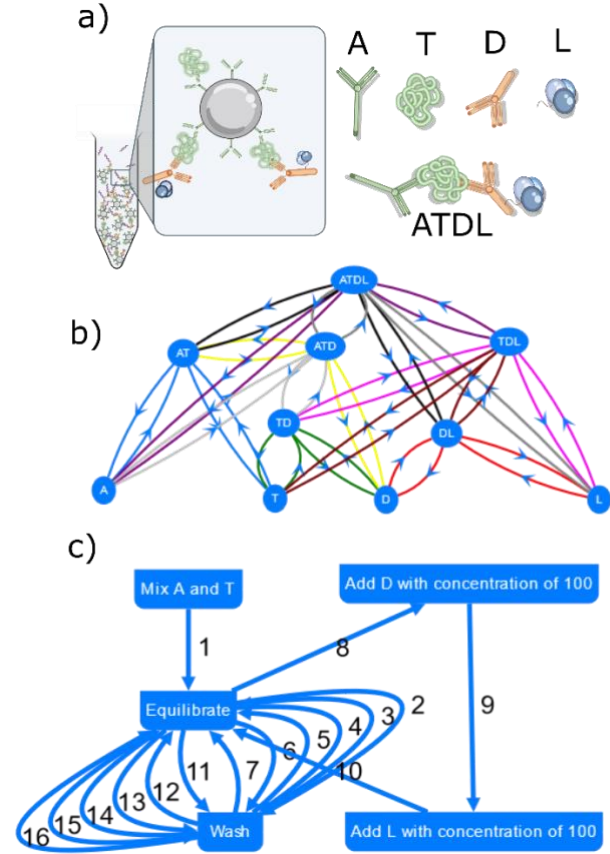


Figure 1: a) Schematic diagram of a typical 4-component bead-based digital immunoassay, in which a capture antibody A bound to a paramagnetic bead capture targets T from solution, which are then labelled with a combination of a detector antibody D and a label L that is eventually used for downstream detection. On the way to assembly of the full complex all possible sub-permutations of 2- and 3-components will form, leading to a total of 10 possible interacting species in solution. b) Network diagram of interactions between the four primary components of a bead-based digital immunoassay and complexes thereof. Paired binding pathways are color-coded to highlight interaction partners in the network. c) A typical assay workflow. Assay begins at the “Mix” step and follows the arrows in the labelled order.

Note that we are not explicitly simulating beads in this case, but instead simulating the capture antibodies as though they are uniformly distributed through the reaction volume. This is a subtlety that is very important: when a bead is coated with N capture antibodies, to first order it acts as a single antibody with a k_{on} is that N times larger than the k_{on} for a single antibody alone while k_{off} remains unchanged. In the case where many of those binding sites are occupied, the effective *on*-rate reduces proportionately, that is, a bead with N binding sites of which n are occupied has an effective *on*-rate of $(N - n)k_{on}$ with respect to the next binding event, which means that it is not

possible to simulate these beads directly unless we consider the time-dependent on-rate that results. Because these beads are spread through the 3-dimensional volume of the reaction chamber, however, it is reasonable from a simulation standpoint to ignore those spatial correlations and to simply simulate the capture antibodies as though they were freely diffusing on their own, which circumvents the need to account for partial binding affecting the *on*-rate of the beads themselves. This is reasonable if the ratio of capture antibody to target molecule is large and steric effects are negligible, which will always be true for any real assay attempting to measure the concentration of a low-abundance target. As an example, in the SiMoA platform, the ratio of target to bead is <1 for a digital readout while each bead contains hundreds of thousands of antibodies.^{5,11} Nevertheless, care should be taken when using this model to simulate a situation in which target concentrations are comparable to the capture antibody concentrations, as in this case these steric effects may be important and this model would be expected to overestimate binding.

A typical paramagnetic microbead-based digital assay workflow generally follows the steps below. This workflow is shown in a flowchart in Figure 1c.

1. Capture-antibody coated beads are mixed with the target and allowed to equilibrate.
2. Beads are immobilized and washed to remove unbound target and any nonspecific background molecules. Note that some captured target will fall off during and after this step, although this is minimized by the enhanced k_{on} as long as the number of capture antibodies is high relative to the number of targets.
3. Beads are resuspended and detection antibodies and labelling enzymes are added and allowed to bind and label the captured targets remaining on the beads.
4. Beads are immobilized and washed to remove unbound detector and label. Note that some captured target-label complexes, as well as labels themselves, will fall off during and after this step.
5. Post-processing done to count the fraction of labelled beads.

As an illustration of the capabilities of this software, we first approximate a bead-based digital assay with nanopore electrical readout which we conducted recently.¹⁵ Unfortunately, due to the fact that detailed *on*- and *off*- rates are not available for the capture antibody and detector antibody pairings in that assay, we must make some approximations. To do so while still gaining useful physical insights, we normalize parameters by the expected behavior of the capture antibody pairing and take generic binding constants for all interactions, setting *on*- and *off*-rates for all components to 1 (arbitrary units), except for the detector-label *on*- and *off*-rates, which are set to 10 and 0.01 to approximate the much stronger biotin-streptavidin interaction used to bind the label to the detector. Target concentration is set to 1, which can be thought of as being equal to K_D for the target-capture antibody pairing. Capture antibody concentration is set to 10^3 to simulate an extreme excess, while detector antibody and label, when added, are at a concentration of 10^2 . In the

actual assay, the excess in both cases was even more by at least an order of magnitude, but we have found in simulation that such disparity in concentrations can lead to numerical instability or very long runtimes, while not actually changing the results significantly. Wash steps are modelled by setting the free concentration of all species that are not bound to a bead (in this case, that do not contain an A) to zero, though the general framework allows for the effects of imperfect washing to be simulated as well by setting the wash efficiency to a number between 0 and 1 for all components. Three washes are performed after allowing the initial equilibration of A and T and after adding D and L. The system is propagated to equilibrium between each step. Figure 1c shows the full workflow being simulated.

Figure 2 presents the time evolution of the relative concentrations of the various interacting species in such a typical 4-component bead-based sandwich immunoassay. Figure 2a is a zoom into step 1, showing the time-evolution of the capture of targets by antibody-coated beads. The target T starts at a normalized concentration of 1 and antibody-coated beads, A, at 10^3 . Equilibration of AT complexes happens very fast due to the excess of capture antibodies, and there is essentially no loss of target to the wash steps since any targets that fall off are immediately recaptured by a nearby capture antibody. Figure 2b then shows the addition of the detector antibody, D, and label, L, and the time-evolution of all complexes that are not removed during the washes (anything containing an A, i.e., stuck to the bead). Concentrations in this part of the assay respond much more strongly to washes, and equilibrium takes orders of magnitude longer to establish. Figure 2c shows the total amount of T and L in the system, demonstrating the loss of signal that occurs as a function of wash timing. While essentially no target is lost throughout the assay, the amount of label available to indicate its initial presence in the sample is highly dependent on the number of wash steps following the addition of the detector antibody, D, and label, L.

From the results of Figure 2 a few things are immediately apparent. The first is that not all wash steps are equal: very little target is lost in the first round of wash steps, while a majority of detector-label complex is lost in the second round. While it is possible to calibrate the assay to account for these losses, the signal will get exponentially weaker in the number of wash steps that occur after the label is added, implying that wash steps should be front-loaded to the extent possible, and that minimal washing is desirable after adding the label. The much longer equilibration time required after a wash at this stage is also good reason to limit post-labelling washes to one if possible. The reason for the asymmetry is simple. Due to the enormous excess of capture antibodies, any targets that fall off almost immediately rebind, whereas there are very few sites for the detectors to attach, meaning that an unbinding event is usually final at this stage. It is interesting to note that most of the losses occur on longer timescales, and that immediately after a wash everything remains bound, approaching equilibrium via a stretched exponential process. This confirms the intuitive understanding that the timing and duration of wash steps is a critical consideration in designing an effective digital immunoassay workflow.

It is interesting to note that in Figure 2c that the full ATDL complex makes up only a minority of the total L available in the system at any given time, with most of the L free-floating and not attached to an antibody once equilibrium is established. This is of significant consequence for downstream readout mechanisms. In the SiMoA assay model, only ATDL complexes give rise to downstream signal detection, since beads need to be confined to microwells for optical readout, whereas in the nanopore assay, all sources of L left over at the end of the washes will contribute, whether still bound or not on the bead. Care must therefore be taken when constructing calibration curves with the downstream detector to ensure that the source of signal, and losses thereto, are properly accounted for in the upstream assay design. Alternatively, one can design a wash schedule to occur before equilibrium is established, in which case wash timing must be strictly adhered to when comparing a calibration curve to a full assay run.

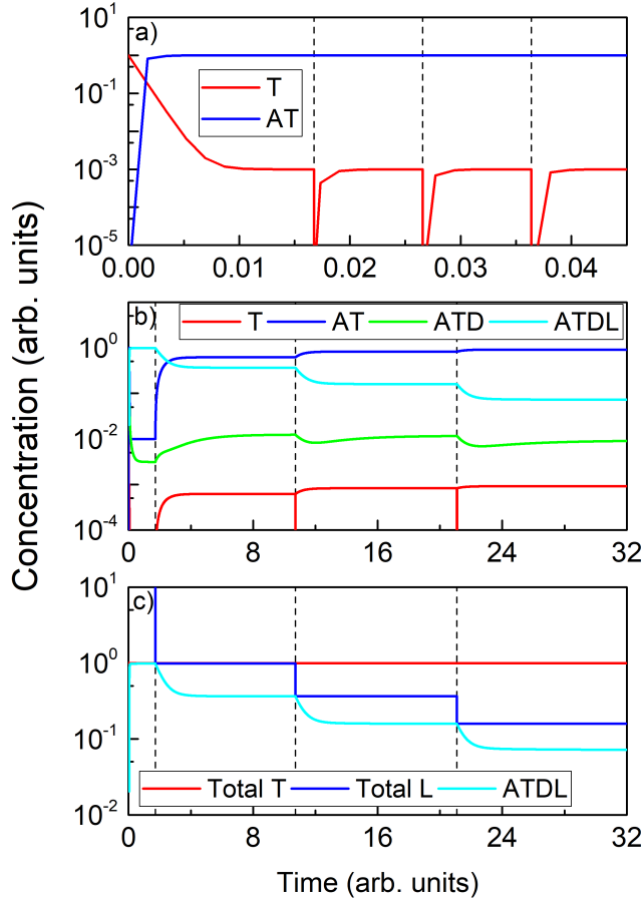


Figure 2: Time evolution of the relative concentrations of the various interacting species in a typical 4-component bead-based sandwich immunoassay. a) fast equilibration of the target molecule T, with an initial concentration of 1, in the presence of an extreme excess of capture antibody A (1000x), forming AT complexes, close to 1 (i.e., ~99.99% of targets are captured very rapidly). b) Time-evolution of complexed species that do not get washed away during wash steps. c) Time evolution of the total target T and total label L available in the system, showing a strong loss of signal that occurs late in the workflow. Vertical dashed lines indicate wash timing.

DNA Nanostructure Binding. We next consider a different binding model that we recently explored experimentally:^{15,21} the

binding together of two star-shaped DNA nanostructures using a linker strand into a dumbbell shape, as shown in Figure 3a. We simulate this using a two-component system initially – stars and linkers. DNA stars have a single stranded region at the end of a double stranded tail that is complementary to half as of ssDNA linker. Once a linker binds to one star, it can bind with an additional star to form the dumbbell. We estimate $k_{on} \sim 10^6 s^{-1} M^{-1}$,²³ and assume that binding is irreversible, so that all *off*-rates are 0. As we will show, this simple model fails to fully capture the experimental results, and we complexify the model from this simple starting point, considering the effect first of misassembled DNA star structures that cannot properly bind as well as stars that can bind into a pseudo-dumbbell in the absence of a linker strand, and finally including the presence of an additional population of misformed stars that can bind into duplexes with a slower *on*-rate. These three cases and comparison of the resulting simulations to experiment are shown in Figure 3.

In its simplest form (Figures 3a and 3b), one intuitively expects that the fraction of dumbbells f formed should be equal to the ratio of linkers to stars x (assuming equal concentrations of both star-halves), or the inverse of that ratio, whichever is smaller, that is:

$$f = \min(x, x^{-1}) \quad (10)$$

This is a consequence of the irreversibility of the binding, since when there is an excess of linker strands the star nanostructures will get capped by the linker strand and be unable to bind further to another star, since two linker strands cannot bind together. This supports the notion that we previously presented that any practical application of these schemes must operate in the regime where linker strands are the limiting reagent, which is practically the case for any real assay.

In the experimental case, shown as the black curve in Figure 3b, Equation 10 only holds true for values of $x \ll 1$. The peak value at $x = 1$, is smaller experimentally than the theoretical prediction. Below $f \sim 10^{-2}$ the dumbbell fraction reaches a minimum (noise floor) that persists even in the absence of linker strands. Finally, there exists an asymmetry showing more binding than expected for $x > 1$. We hypothesized that the first discrepancy is likely due to a fraction of misassembled stars that cannot bind properly, while the second one is likely due to misassemblies that occur in such a way as to allow the stars to bind together in the absence of a linker. The simulation tool allows us to validate these hypotheses. If we include both types of misassembled products in the simulation, we indeed recover both the reduced maximum value of f and the false positive behavior at both ends of the spectrum, as shown in Figure 3b.

It is interesting to note the discrepancy between even the modified simulation and experiment for $x > 1$, where simulation still underestimates the binding, and an asymmetry is present in the experimental results that is not predicted by Equation 10. A hypothesis to explain this discrepancy is that what we are modelling as non-binding misassemblies instead bind less strongly with linkers, and when linker strands are present in extreme excess, these weakly binding species begin to matter for the kinetics. This scenario could arise, for example, if the single-stranded tail meant to interact with the linker

instead weakly binds to a misassembled single-stranded arm through some weak base-pairing that would then require a strand displacement reaction for the linker to bind properly. Were this the case, one would expect that strand displacement only to happen in the case where there is an excess of linkers leftover after binding the unaffected probes.

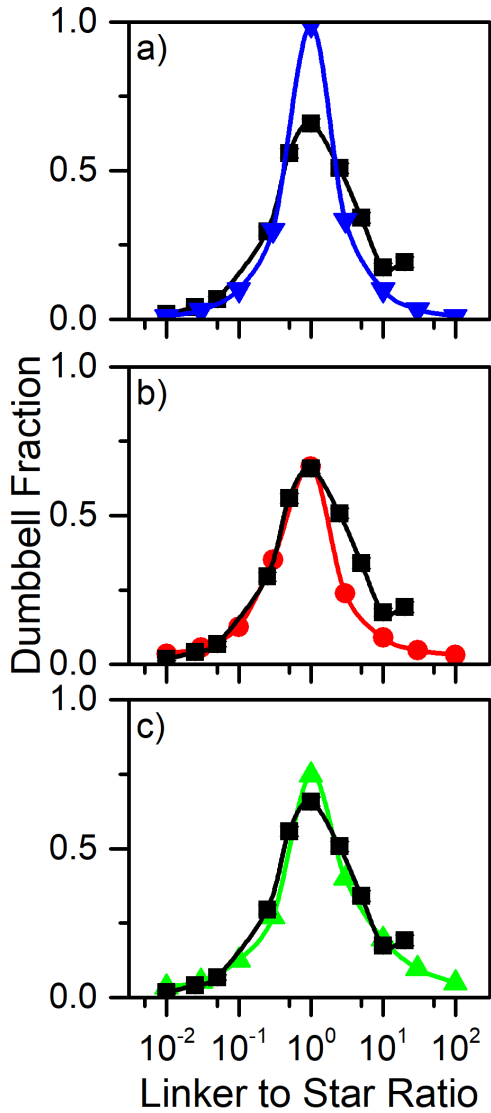


Figure 3: a) A schematic illustration of the two-step binding scheme of pairs of probes into dumbbells. b) a comparison between an idealized simulation (blue triangles) and experimental results from He *et al*¹⁵ (black squares). c) An illustration of the effects of non-binding misassembled probes and of the effects of probes that can bind in the absence of a linker strand. d) An illustration of the asymmetry introduced when a subset of probes has weaker and competitive binding kinetics with the linker. For the experimental data, the linker strand concentration ranges from 200 pM to 400 nM and shooting star probes is fixed at 20 nM, with ~1100 single-molecule events at each concentration.

To test this hypothesis, we simulated this as well, using an *on*-rate for the strand displacement component that is 10x smaller than the *on*-rate for the properly assembled DNA stars. The introduction of these binding elements reproduces the

asymmetry observed experimentally, as the linkers will bind the misassemblies in significant quantities only in the case where there is an excess beyond that required to saturate the properly assembled probe molecules. This agreement can be seen in Figure 3c.

Note that the total concentration of dumbbells remains unchanged when introducing misassemblies, but is broken up between proper assemblies and misassemblies, and that the reported dumbbell fraction included in the numerator the sum of all structures that would look like a dumbbell to a nanopore (two stars bound together by a linker strand), and in the denominator the sum of the concentrations of all structures that would look like stars to the nanopore.

Because the actual rate constants are unknown and simply chosen to show qualitative effects of having different kinds of misassemblies present, the concentration breakdown is arbitrarily chosen to match the experimental data. It should be carefully noted at this point that the fact that a model accurately reproduces experimental behavior is not definitive proof that that model is an accurate representation of the physics, nor are the rate constants or concentrations chosen expected to be an accurate reproduction of the real breakdown of misassemblies. This tool is not meant to prove that a system of interaction is the underlying physical system. Rather, it allows rapid exploration of the downstream expectation if certain interactions are included so that physical intuition and hypotheses can be rapidly tested for consistency without needing to invest upfront experimental time and resources.

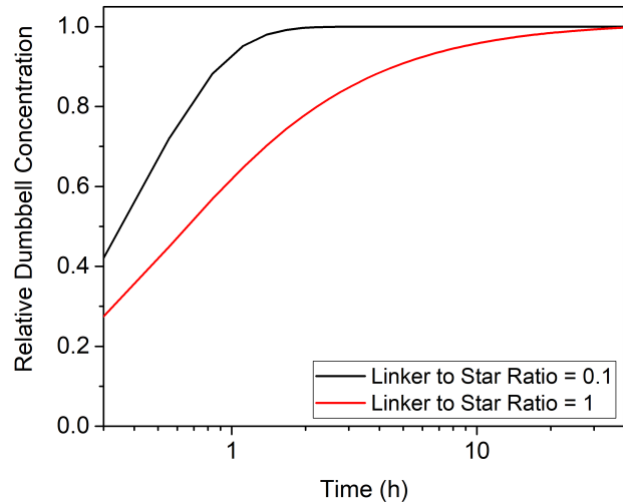


Figure 4: A comparison of time-to-equilibrium for two different ssDNA linker to dsDNA stars ratios, demonstrating that equilibrium takes much longer to establish when binary components are present in equal concentration. Note that time in log-scaled here to show the different more clearly.

The second issue that arises in such first-order binding systems is that the two-step binding required to make a full DNA nanostructure greatly extends the time required to reach equilibrium when $x \sim 1$, experimentally taking more than 24 hours at parity (see SI He *et al*¹⁵). Figure 4 presents a comparison of the time response of the concentration of dumbbells normalized to the equilibrium concentration ratios of $x = 1$ and $x = 0.1$. Equilibrium takes much longer to establish

at parity (red curve), even though the concentration of DNA stars in both cases is the same.

This can be understood, under the assumption of irreversible binding, by the fact that at parity the concentration of available reagents to bind is depleted as the reaction progresses, leading to progressively slower kinetics. The need for two binding events to form the positive molecular signal comes at the cost of increased assay time. Which is clearly seen in Figure 4, and suggests that multi-step bindings are to be avoided to minimize assay time when required reagents are present in equal concentrations. Future digital assays with nanopore readout should take this into consideration

Conclusions. We have presented a computational tool that allows for simulation of arbitrarily complex receptor-ligand or DNA hybridization first-order kinetics-based assay workflows from start to finish without requiring any programming knowledge on the part of the user. We have demonstrated its utility for providing insights into assay performance and validating hypotheses, through two case studies. We compared our experimental results with those of the simulation tool, finding good agreement. A detailed description of the user interface along with a basic user guide is included in the Supporting Information. The webapp is freely available at immunoassay.herokuapp.com.

Many ligand-receptor pairs do not have separate *on*- and *off*-rate values listed in the literature, and only have available the equilibrium constant for the reaction, that is $K_d = \frac{k_{off}}{k_{on}}$. In this case an educated guess must be made as to how this quantity splits into *on*- and *off*-rates individually, but by varying these parameters and simulating over the possible splits, the simulation tool can provide some insight into the effect of these assumptions and allow for matching to experimental results. In a simple two-component reaction the equilibrium state of the system depends only on K_d ; for more complex interacting systems, or if measuring concentrations at intermediate times prior to equilibrium where the system is still in a transient state, the individual rates matter. Even in the absence of exact *on*- and *off*-rate values to provide quantitative results, the tool offers a method by which to validate assay designs generally, giving clear insight into the effect on downstream signal of wash steps and nonspecific binding of background molecules, allowing full assay workflows to be validated in minutes prior to conducting expensive and time-consuming experimental work that can take weeks. This application of our simulation tool was demonstrated in our first case study. It is also a tool with which to test ideas relating to experimental workflow failure, making it simple to test hypotheses about any discrepancies between prediction and experiment by defining potential unintended component interactions, as we demonstrated in our second case study.

Finally, the results presented so far have provided valuable insight into bead-based digital immunoassay design generally: wash steps should be conducted as early in the assay workflow as practicable and must be carefully timed to be consistent between calibration and experiment if not allowed to equilibrate between washes.

AUTHOR INFORMATION

Corresponding Author

*tcossa@uottawa.ca

Author Contributions

KB wrote the first draft of the manuscript and wrote the backend of the software used to perform the simulations and subsequent analysis. MYB wrote the frontend, performed web integration, and designed and implemented visualization tools. LH performed bead-based nanopore immunoassay experiments. All authors contributed edits to the final version of the manuscript.

Notes

KB and VTC declare competing financial interest as cofounders of Northern Nanopore Instruments Inc. a for-profit company which sells solid-state nanopore-related hardware and software.

ACKNOWLEDGMENT

The authors would like to acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC, grant #CRDPJ 530554-18). KB thanks Sasha Rogers Van Katwyk for many fruitful discussions around modelling techniques and for making clear that binary assay kinetics modelling and local contagious disease modelling are fundamentally the same thing.

REFERENCES

- (1) Engvall, E.; Perlmann, P. Enzyme-Linked Immunosorbent Assay (ELISA). Quantitative Assay of Immunoglobulin G. *Immunochemistry* **1971**, *8* (9), 871–874. [https://doi.org/10.1016/0019-2791\(71\)90454-x](https://doi.org/10.1016/0019-2791(71)90454-x).
- (2) Gosling, J. P. A Decade of Development in Immunoassay Methodology. *Clin. Chem.* **1990**, *36* (8 Pt 1), 1408–1427.
- (3) Zhang, H.; Zhao, Q.; Li, X. F.; Le, X. C. Ultrasensitive Assays for Proteins. *Analyst*. 2007. <https://doi.org/10.1039/b704256f>.
- (4) Giljohann, D. A.; Mirkin, C. A. Drivers of Biodiagnostic Development. *Nature* **2009**, *462* (7272), 461–464. <https://doi.org/10.1038/nature08605>.
- (5) Walt, D. R. Optical Methods for Single Molecule Detection and Analysis. *Anal. Chem.* **2013**, *85* (3), 1258–1263. <https://doi.org/10.1021/ac3027178>.
- (6) Rissin, D. M.; Kan, C. W.; Campbell, T. G.; Howes, S. C.; Fournier, D. R.; Song, L.; Piech, T.; Patel, P. P.; Chang, L.; Rivnak, A. J.; Ferrell, E. P.; Randall, J. D.; Provuncher, G. K.; Walt, D. R.; Duffy, D. C. Single-Molecule Enzyme-Linked Immunosorbent Assay Detects Serum Proteins at Subfemtomolar Concentrations. *Nat. Biotechnol.* **2010**, *28* (6), 595–599. <https://doi.org/10.1038/nbt.1641>.
- (7) Connie Wu; Garden, P. M.; Walt, D. R. Ultrasensitive Detection of Attomolar Protein Concentrations by Dropcast Single Molecule Assays. *J Am Chem Soc* **2020**, *142* (28), 12314–12323. <https://doi.org/doi:10.1021/jacs.0c04331>.
- (8) Kelley, S. O.; Mirkin, C. A.; Walt, D. R.; Ismagilov, R. F.; Toner, M.; Sargent, E. H. Advancing the Speed, Sensitivity and Accuracy of Biomolecular Detection Using Multi-Length-Scale Engineering. *Nat. Nanotechnol.* **2014**, *9* (12), 969–980. <https://doi.org/10.1038/nnano.2014.261>.
- (9) Mao, C.-P.; Wang, S.-C.; Su, Y.-P.; Tseng, S.-H.; He, L.; Wu, A. A.; Roden, R. B. S.; Xiao, J.; Hung, C.-F. Protein Detection in Blood with Single-Molecule Imaging. *Sci. Adv.* **2021**, *7* (33). <https://doi.org/10.1126/sciadv.abg6522>.
- (10) Wilson, D. H.; Rissin, D. M.; Kan, C. W.; Fournier, D. R.; Piech, T.; Campbell, T. G.; Meyer, R. E.; Fishburn, M. W.; Cabrera, C.; Patel, P. P.; Frew, E.; Chen, Y.; Chang, L.; Ferrell, E. P.; von Einem, V.; McGuigan, W.; Reinhardt, M.; Sayer, H.; Vielsack, C.; Duffy, D. C. The Simoa HD-1 Analyzer: A Novel Fully Automated Digital Immunoassay Analyzer with Single-Molecule Sensitivity and Multiplexing. *J. Lab. Autom.* **2016**, *21* (4), 533–547. <https://doi.org/10.1177/2211068215589580>.
- (11) Chang, L.; Rissin, D. M.; Fournier, D. R.; Piech, T.; Patel, P. P.; Wilson, D. H.; Duffy, D. C. Single Molecule Enzyme-Linked

- Immunosorbent Assays: Theoretical Considerations. *J. Immunol. Methods* **2012**, *378* (1–2), 102–115. <https://doi.org/10.1016/j.jim.2012.02.011>.
- (12) Morin, T. J.; Mckenna, W. L.; Shropshire, T. D.; Wride, D. A.; Deschamps, J. D.; Liu, X.; Stamm, R.; Wang, H.; Dunbar, W. B. A Handheld Platform for Target Protein Detection and Quantification Using Disposable Nanopore Strips. *Sci. Rep.* **2018**, *8* (January), 1–12. <https://doi.org/10.1038/s41598-018-33086-7>.
- (13) Shave, S.; Chen, Y.-K.; Pham, N. T.; Auer, M. PyBindingCurve, Simulation, and Curve Fitting to Complex Binding Systems at Equilibrium. *J. Chem. Inf. Model.* **2021**, *61* (6), 2911–2915. <https://doi.org/10.1021/acs.jcim.1c00216>.
- (14) Bruce, N. J.; Ganotra, G. K.; Richter, S.; Wade, R. C. KBbox: A Toolbox of Computational Methods for Studying the Kinetics of Molecular Binding. *J. Chem. Inf. Model.* **2019**, *59* (9), 3630–3634. <https://doi.org/10.1021/acs.jcim.9b00485>.
- (15) Liqun He; Tessier, D. R.; Briggs, K.; Tsangaris, M.; Charron, M.; McConnell, E. M.; Lomovtsev, D.; Tabard-Cossa, V. Digital Immunoassay for Biomarker Concentration Quantification Using Solid-State Nanopores. *Nat. Commun.* **2021**, *Accepted*.
- (16) Beamish, E.; Tabard-Cossa, V.; Godin, M. Digital Counting of Nucleic Acid Targets Using Solid-State Nanopores. *Nanoscale* **2020**. <https://doi.org/10.1039/d0nr03878d>.
- (17) Kwok, H.; Briggs, K.; Tabard-Cossa, V. Nanopore Fabrication by Controlled Dielectric Breakdown. *PLoS One* **2014**, *9* (3), e92880. <https://doi.org/10.1371/journal.pone.0092880>.
- (18) Briggs, K.; Charron, M.; Kwok, H.; Le, T.; Chahal, S.; Bustamante, J.; Waugh, M.; Tabard-Cossa, V. Kinetics of Nanopore Fabrication during Controlled Breakdown of Dielectric Membranes in Solution. *Nanotechnology* **2015**, *26* (8), 084004. <https://doi.org/10.1088/0957-4484/26/8/084004>.
- (19) Waugh, M.; Briggs, K.; Gunn, D.; Gibeault, M.; King, S.; Ingram, Q.; Jimenez, Aura Melissa Berryman, S.; Lomovtsev, Dmytro Andrzejewski, Lukasz Tabard-Cossa, V. Solid-State Nanopore Fabrication by Automated Controlled Breakdown. *Nat. Protoc.* **2019**, *15*, 122–143. <https://doi.org/10.1038/s41596-019-0255-2>.
- (20) Leung, C.; Briggs, K.; Laberge, M.-P.; Peng, S.; Waugh, M.; Tabard-Cossa, V. Mechanisms of Solid-State Nanopore Enlargement under Electrical Stress. *Nanotechnology* **2020**, *31* (44). <https://doi.org/10.1088/1361-6528/aba86e>.
- (21) He, L.; Karau, P.; Tabard-Cossa, V. Fast Capture and Multiplexed Detection of Short Multi-Arm DNA Stars in Solid-State Nanopores. *Nanoscale* **2019**, *11* (35), 16342–16350. <https://doi.org/10.1039/c9nr04566j>.
- (22) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes: The Art of Scientific Computing 3rd Edition*; Cambridge University Press, 2007.
- (23) Zhang, J. X.; Fang, J. Z.; Duan, W.; Wu, L. R.; Zhang, A. W.; Dalchau, N.; Yordanov, B.; Petersen, R.; Phillips, A.; Zhang, D. Y. Predicting DNA Hybridization Kinetics from Sequence. *Nat. Chem.* **2018**, *10* (1), 91–98. <https://doi.org/10.1038/nchem.2877>.