Published online 13 September 2022 *Nucleic Acids Research, 2022, Vol. 50, No. 17* **9647–9662**
https://doi.org/10.1093/nar/gkac784

Hybridization kinetics of out-of-equilibrium mixtures of short RNA oligonucleotides

Marco Todisco [1,2] and Jack W. Szostak[1,2,3,*]

[1]Howard Hughes Medical Institute, Department of Molecular Biology and Center for Computational and Integrative Biology, Massachusetts General Hospital, 185 Cambridge Street, Boston, MA 02114, USA, [2]Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA and [3]Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA

Received June 09, 2022; Revised August 23, 2022; Editorial Decision August 26, 2022; Accepted September 08, 2022

ABSTRACT

Hybridization and strand displacement kinetics determine the evolution of the base paired configurations of mixtures of oligonucleotides over time. Although much attention has been focused on the thermodynamics of DNA and RNA base pairing in the scientific literature, much less work has been done on the time dependence of interactions involving multiple strands, especially in RNA. Here we provide a study of oligoribonucleotide interaction kinetics and show that it is possible to calculate the association, dissociation and strand displacement rates displayed by short oligonucleotides (5nt–12nt) that exhibit no expected secondary structure as simple functions of oligonucleotide length, CG content, ΔG of hybridization and ΔG of toehold binding. We then show that the resultant calculated kinetic parameters are consistent with the experimentally observed time dependent changes in concentrations of the different species present in mixtures of multiple competing RNA strands. We show that by changing the mixture composition, it is possible to create and tune kinetic traps that extend by orders of magnitude the typical sub-second hybridization timescale of two complementary oligonucleotides. We suggest that the slow equilibration of complex oligonucleotide mixtures may have facilitated the nonenzymatic replication of RNA during the origin of life.

INTRODUCTION

Characterizing the behavior of mixtures of nucleic acids is important in a wide variety of fields, including gene silencing (1–3), primer design for routine PCR experiments, DNA nanotechnology (4,5), nucleic acid-based computing (6), the development of novel materials (7–9), and the study of RNA World scenarios relevant to understanding the origins of life (10–13). The pillar on which all these studies have been built is the nearest-neighbor thermodynamic model for the stability of base paired nucleic acid duplexes (14,15). This model accurately predicts the change in free energy (ΔG°) when two oligonucleotides anneal, at least in cases where the process can be approximated as a two-state transition and assuming that the oligonucleotides have negligible secondary structure and self-interactions.

Although the nearest-neighbor model provides a good description of the equilibrium state of the system, a solution of many different oligonucleotides may require a long time to equilibrate upon mixing, because the system can become trapped in local energy minima. In such cases escaping from metastable states may require thermal annealing to reach the global minimum. Depending on the system, such out-of-equilibrium states can either constitute a strong disadvantage, causing the formation of unwanted duplexes whose concentration varies with time, or it can be a valuable source of complexes with single-stranded overhangs or gaps that can behave as substrates for desired physical processes (16,17), or for chemical reactions such as template copying or loop closing under the proper activating conditions (12,18–20). While the change in free energy can ultimately be calculated with a reasonable degree of accuracy from oligonucleotide sequences and ionic conditions, a more rigorous and extensive study of hybridization kinetics together with modelling is required to better understand the out-of-equilibrium behavior of mixtures of RNA oligonucleotides.

When two complementary oligonucleotides are mixed in solution, they rapidly hybridize to produce a double helix, annealing *via* a two-step process involving nucleation by the formation of several consecutive paired bases, followed by a rapid zippering step. Because the nucleation step is generally assumed to be rate-limiting, duplex formation in a simple case with no secondary structures (21,22) can be modeled as a bimolecular reaction with an effective rate k_{on} (see Table 1 and Supplementary Table S1 for list of names and definitions), dominated by the time necessary for the two strands to collide and then form a productive nucleation region that will zipper (23–26). The kinetics of oligonucleotide

*To whom correspondence should be addressed. Tel: +1 617 726 5102; Fax: +1 617 643 3328; Email: szostak@molbio.mgh.harvard.edu

binding has traditionally been measured using techniques such as following the FRET signal between fluorescently labeled oligonucleotides (1,27,28), following changes in UV or IR absorbance (29–33) or using specific sensors (34–37) that rely on DNA binding to a functionalized surface.

With the aim of better understanding nucleic acid hybridization, a series of modeling approaches have been explored in the past, from approximate analytical treatments (31,38) to Markov-chain stochastic simulations (39) and molecular dynamics coarse-grained simulations (23,40). Recently, two noteworthy approaches have been established for predicting the hybridization kinetics of nucleic acids: (i) Zhang *et al.* published a tool based on a complex Weighted Neighbor Voting algorithm built on a database of 36nt long DNA oligonucleotides (41), while (ii) Rejali and colleagues attempted to deconstruct the hybridization rate of a set of short DNA strands with a NN parameterization. Unfortunately, this latter approach showed a limited reliability and, to overcome this issue, the authors proposed a simplified predictive function based on the CG content alone of the analyzed strands. One caveat of these two recent predictive approaches is that they neglect the potential role of oligonucleotide length in association kinetics. The role of length remains controversial in the literature: early studies on long DNA molecules have clearly shown a positive dependence of k_{on} on the length of the oligonucleotides (38,42). However, the behavior of short oligonucleotides is still a matter of debate, with recent studies claiming to show hybridization rates being either positively dependent, independent or negatively dependent on length (23,25,27,29,31). To further complicate the picture of binding kinetics, studies on DNA and RNA by Cisse *et al.* have shown that a minimum of seven consecutive complementary nucleotides are necessary for fast annealing, with a significant drop in rate for shorter stretches (1). This observation suggests that the minimum nucleation site could be larger than previously expected and could thus dramatically influence k_{on} for the short DNA and RNA oligonucleotides that are the primary focus of our work.

Following their hybridization, two oligonucleotides are not irreversibly bound; instead, they will dissociate with a characteristic dissociation rate k_{off}. While literature k_{on} values for short DNA and RNA molecules do not typically vary by more than an order of magnitude across different oligonucleotide compositions and lengths, reported duplex lifetimes show a strong dependence on the length and thermodynamic stability of the oligonucleotide duplexes, resulting in k_{off} values spanning many orders of magnitude (27,31,43).

When more than two complementary sequences are present in solution, individual association and dissociation rates are not sufficient to provide a complete description of the time-dependent evolution of the system. An important phenomenon that must be accounted for is toehold-mediated strand displacement. This occurs when one strand of a duplex is displaced by an incoming invading strand, a phenomenon that is widely exploited in DNA nanotechnology to develop complex circuits (44–46). More recently toehold-mediated strand displacement has been applied in the field of the origin of life to enable enzyme-free copying of duplex RNA (47). A seminal study by Zhang and Winfree (48) has clarified the details of this process in DNA and shown how the magnitude of the strand displacement rate depends on the binding of the incoming oligonucleotide to an unpaired region of the template, the so-called toehold. Only recently have the kinetics of toehold-mediated strand displacement for RNA been computationally (49) and experimentally investigated, using an approach analogous to that of the Winfree group by relying on a fluorescent reporter to monitor strand displacement over a region longer than 20 bp (50). However, further studies are needed to provide a deeper understanding of this process in shorter oligonucleotides and establish a reliable and comprehensive model for the prediction of rates of toehold-mediated strand displacement in RNA oligonucleotides.

Together, k_{on}, k_{off} and toehold-mediated strand displacement rates should enable the description of the evolution of a mixture of oligonucleotides from their mixing to their equilibrium state, giving valuable information on the state of the system at intermediate times. In this work we present a study of RNA annealing kinetics and provide a framework for the prediction of the time-dependent evolution of mixtures of short oligoribonucleotides with no secondary structures.

MATERIALS AND METHODS

General

Controlled pore glass (CPG) columns, phosphoramidites and reagents for oligonucleotide synthesis and purification were from Glen Research. Reagents for cleavage and deprotection were from Sigma-Aldrich. Buffers were prepared from 1 M Tris stock solutions (Invitrogen) and pH-adjusted using HCl. Sodium chloride and magnesium chloride hexahydrate were from Sigma-Aldrich. Unless otherwise stated all measurements were performed in 200 mM Tris–HCl pH 8 and 100 mM $MgCl_2$. Because the literature standard for nucleic acids is 1 M NaCl, we performed a subset of experiments in 5 mM Tris–HCl 1 M NaCl pH 7 to compare the two conditions and no significant difference was noted, and so we consider these two conditions as equivalent (see Supplementary Figures S1 and S2).

Solid phase synthesis of oligoribonucleotides

RNA oligonucleotides were synthesized in-house on an Expedite 8909 by solid-phase synthesis using the DMT-on protocol. Synthesized oligonucleotides were cleaved from the RNA-CPG columns with AMA (1.2 ml, 1:1 mixture of 40% methylamine solution and 28% NH_4OH) for 30 min. The solution containing the oligonucleotides was transferred from the column to a screw-cap vial and incubated at 65°C for 20 min for deprotection. The resulting solution was dried for 30 min at 40°C in a speed-vac and then lyophilized overnight. To remove the TBDMS protecting groups the resultant powder was resuspended in DMSO (115 ul), TEA (60 ul) and TEA·3HF (75 ul) and incubated at 65°C for 2.5 h. Oligonucleotides were purified from the mixture and the DMT group removed using GlenPak columns. List of all sequences used is available in Supplementary Table S2.

Oligonucleotide concentrations were determined spectrophotometrically using a NanoDrop 2000 (Thermo Sci-

Table 1. Table of names

Name	Definition
k_{on}	Hybridization rate of two oligonucleotides
k_{off}	Detaching rate of two oligonucleotides
K_D	Dissociation constant
K_A	Association constant
k_{bi}	Bimolecular collision rate for the formation of a contact between two oligonucleotides
k_{uni}	Unimolecular rate for the formation of a new base pair in a zippering process
k_{displ}	Bimolecular rate of toehold mediated strand displacement.
k_s	Unimolecular rate of branch migration after the toehold is bound.
k_{mig}	Unimolecular rate for the single step of branch migration during a strand displacement.

entific) using the extinction coefficients computed according to Tataurov *et al.* (51).

We monitored the extent of hybridization of one oligonucleotide by measuring the fluorescent emission of 2-aminopurine (2Ap), a fluorescent adenine base analogue whose quantum yield diminishes when in a double helix. In the absence of published values for the extinction coefficients of 2Ap-containing RNA strands, we found that the extinction coefficients provided for 2Ap-DNA by Xu and Nordlund (52) and by the IDT oligo analyzer tool closely approximate the values of 2Ap-RNA, validated by us through quantitative quenching experiments against complementary strands of known concentration (see Supplementary Figure S3).

Measurement of binding affinity through melting experiments

Although there is a wide body of literature on computing the free energy of binding and thus the affinity of complementary oligonucleotides, we began by verifying the predictions of the nearest neighbor-model in our specific experimental system, in order to exploit the model and the crucial relationship between K_D and binding and dissociation kinetics ($K_D = k_{off}/k_{on}$).

We measured the binding energies of highly stable oligonucleotide duplexes ($\Delta G < -10$ kcal/mol) through melting experiments monitored using either a Jasco FP-8500 Spectrofluorometer equipped with ETC-815 Peltier temperature-controlled cell holder (for fluorescently labeled 2-aminopurine containing oligonucleotides) or an Agilent 3500 UV-Vis Spectrophotometer (for non-fluorescent oligonucleotides). For each pair of oligonucleotides, we measured melting curves using a temperature ramp of 3°C/min. All curves acquired this way have been corrected for baselines at low and high temperature and normalized from zero to one to determine the fraction of unbound oligonucleotides at every temperature. One of the advantages of measuring the thermal stability of a duplex by fluorescence over conventional UV absorbance, is that we can use a complementary non-fluorescent oligonucleotide at a very high concentration while maintaining a high signal to background ratio. Studying the mixture of a 2Ap-containing oligonucleotide (A) and a binding non-fluorescent oligonucleotide (B), having concentrations respectively [A] and [B], with [B] = [A] + Δ, we can track the fluorescent signal resulting from the melting of the duplex and the release of A during a temperature ramp. The normalized amplitude of the fluorescent trace reflects the fraction (f) of unbound fluorescent oligonucleotide [A] over its total concentration [A_{tot}].

We can write the dissociation constant K_D for the binding of A and B as follows:

$$K_D = \frac{[A][B]}{[AB]} \tag{1}$$

We can substitute concentrations, [B] = [A] + Δ and [AB] = [A_{tot}] - [A]:

$$K_D = \frac{[A]([A] + \Delta)}{[A_{tot}] - [A]} \tag{2}$$

And then substitute [A] as $f \cdot [A_{tot}]$ and rearrange:

$$K_D = \frac{(f \cdot [A_{tot}])^2 + f \cdot [A_{tot}]\,\Delta}{[A_{tot}](1 - f)} \tag{3}$$

Using this expression, we can readily calculate K_D at every temperature of our melting curves since [A_{tot}] and Δ are known by preparation. A Van't Hoff plot for K_D versus $1/T$ can then be used to obtain ΔH and ΔS according to the following function:

$$\ln(K_D) = -\frac{\Delta H}{R \cdot T} + \frac{\Delta S}{R} \tag{4}$$

For UV melting experiments, the two oligonucleotides were introduced at equal concentration to maximize the signal to noise ratio. In this condition the mathematical description simplifies and the thermodynamic parameters are extracted from a Van't Hoff plot of $\ln(c_{tot}/4)$ versus $1/T_m$, where $c_{tot}/4$ is the dissociation constant at the T_m and T_m is the melting temperature of the duplex (the temperature at which the fraction of duplexes in solution is equal to 0.5) in Kelvin.

Stopped-Flow measurements of association kinetics

For fluorescence measurements we find a lower bound for useful signal detection of 2-aminopurine at ≈10 nM using a 1 cm path-length quartz cuvette. Because k_{on} for nucleic acids at high ionic strength is typically larger than 10^6 M^{-1} s^{-1}, reactions at accessible concentrations reach equilibrium in a matter of seconds and are thus too fast to be properly studied by hand mixing. The time dependence of binding was therefore measured using a Jasco FP-8500 Spectrofluorometer equipped with the SFS-852T Stopped-Flow accessory, using samples with concentrations in the micromolar range. To convert fluorescent signals to concentrations, all collected data were normalized using a con-

trol experiment where the fluorescent oligonucleotide was mixed only with buffer. Data processed in this way were fit in MATLAB as a simple bimolecular reaction modeled through a set of numerically integrated differential equations. For weakly binding oligonucleotides, k_{off} and k_{on} were globally fit over a series of measurements at different concentrations (see Supplementary Data 4 for detailed description of fitting procedure and Supplementary Table S3 for list of experiments performed).

Measurement of strand displacement

If the binding of an oligonucleotide to a toehold region equilibrates on a time scale much faster than the subsequent strand displacement reaction, which is likely for the short toehold lengths explored in this work, we can approximate the strand displacement process as a simple second order reaction with an effective second order rate constant $k_{displ} = k_s \cdot K_A$, where k_s is the rate for the unimolecular process of branch migration and K_A is the association constant of the toehold with the incoming invading oligonucleotide ($K_A = K_D^{-1}$).

To measure strand displacement, we exploited the properties of 2-aminopurine to track either (i) the release of a bound oligonucleotide bearing the fluorescent moiety as an increase in fluorescence emission over time, (ii) the quenching of a 2-aminopurine nucleotide located in the toehold region of the templating oligonucleotide or (iii) the quenching of 2-aminopurine located in the incoming strand-displacing oligonucleotide. Data collected in this way for mixtures at various concentrations of displacing oligonucleotides were fit as second order irreversible reactions with rate k_{displ}, and a subset of experiments has been repeated by moving 2Ap from the toehold to the duplex region to rule out any potential interference from the fluorescent nucleotide (see Supplementary Table S4). We find that measurements are comparable regardless of 2Ap position, consistent with a negligible effect of 2Ap on the overall RNA thermodynamics and kinetics.

Measurement of the time-dependent evolution of oligonucleotide mixtures

To measure the equilibration process in a mixture of oligonucleotides, we prepared solutions of oligonucleotides competing for shared binding sites. For a typical experiment four sequences A_S, A_L, B_S, B_L of short (S) or long (L) length are mixed. The A and B sequences are complementary, so that every A sequence can interact with every B sequence while self-interactions are negligible at our experimental concentrations. To track the evolution of the system, we strategically placed 2-aminopurine in B_L so that it could be quenched only by A_L. For every experiment two solutions, one containing A_S and A_L and one containing B_S and B_L were mixed and 2-aminopurine fluorescence was monitored over time. When the four oligonucleotides are initially mixed (i.e. at t_0) we observe the fluorescent signal coming from B_L (see Figure 1), but within a few seconds only double helices A_LB_L, A_SB_S, A_LB_S and A_SB_L are present and the residual signal from 2Ap comes only from A_SB_L duplexes leading to partial quenching of the 2Ap signal. Over time the system converges to its energy minimum where only A_LB_L and A_SB_S duplexes are present and all fluorescence is quenched. The rate of fluorescence decay can therefore be treated as a result of the complex interplay of k_{on}, k_{off} and k_{displ}, as the system converges to its energy minimum.

RESULTS

Experimental design

Although fluorescence-based detection of annealing is advantageous because of high sensitivity and easy accessibility, the conjugation of dyes to oligonucleotides can change their properties (53–55). To avoid altering the properties of the oligonucleotides while also retaining the advantages of fluorescence, we use 2-aminopurine (2Ap) as an internal reporter of the state of the system. 2Ap is a fluorescent adenine analogue with an excitation maximum at ≈305 nm and an emission maximum at ≈370 nm. The quantum yield of 2Ap is diminished in a stacked conformation and therefore its fluorescence emission varies according to the extent of oligonucleotide hybridization (56). The substitution of adenine by 2-aminopurine in DNA is reported to have only a small duplex-destabilizing effect, measured as 0.5 ± 0.9 kcal/mol at 30°C in 1 M NaCl (57).

Because this work is based on the use of oligonucleotides containing a single 2-aminopurine residue, we proceeded to test the effect of this substitution in the context of RNA sequences at high salt concentration. By measuring melting temperatures of oligonucleotides incorporating 2Ap, we estimated a destabilization effect of ≈0.2 kcal/mol at 37°C (see Supplementary Figure S1). This effect is therefore negligible at room temperature and is within typical measurement error. In light of this result, we assume that oligonucleotides containing a single 2Ap are a good proxy for the properties of unmodified RNA (58).

Association kinetics of short oligoribonucleotides

Upon mixing of multiple oligonucleotides in solution they will quickly anneal to form a mixture of all possible double helices resulting from base pairing. The initial abundance of these duplexes in a competition scenario will reflect the relative hybridization rates (k_{on}) for any pair of strands. It follows that characterization of the behavior of any mixture of short RNA oligonucleotides requires the measurement and modeling of k_{on}, which we address in this first section, where we aim to provide the tools to readily calculate hybridization kinetics for any given pair of sequences of interest and thereby determine the initial configuration of an arbitrary collection of RNA strands.

The sequences used in this work were designed to minimize secondary structures, and map onto six template sequences with CG contents ranging from 0.0 to 0.8 (see Supplementary Table S1 for a list of all sequences). Due to the high binding energy of CG base pairs, we could not further increase the CG content of our oligonucleotides without creating significant secondary structures. In addition, guanine-rich repetitive structures are known to self-assemble to form G-quadruplexes (59,60). To avoid the formation of structures that would compete with hybridiza-

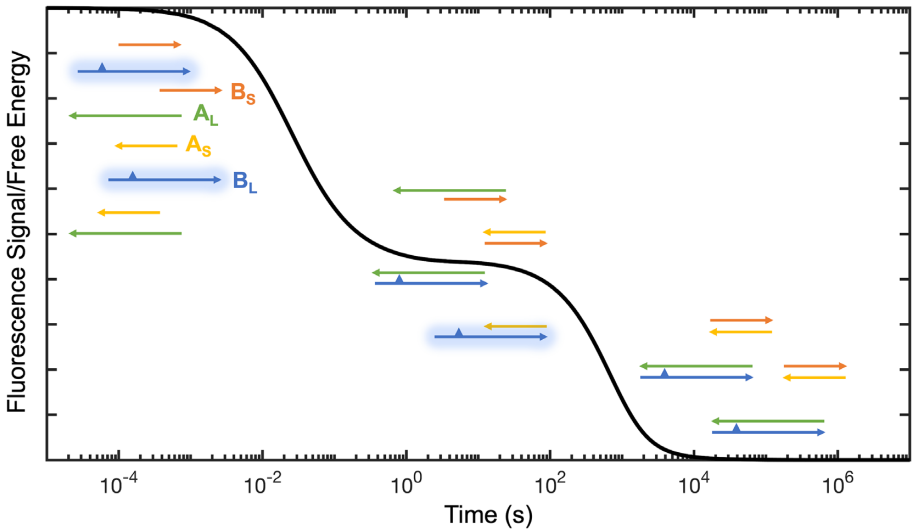

Figure 1. Experimental design for measurement of equilibration time in mixtures of oligonucleotides. At the beginning of the experiment the fluorescent signal of 2Ap (depicted as a triangle) coming from oligonucleotide B_L is maximal, and is progressively quenched as more $A_L B_L$ duplex is formed. A curve depicting a hypothetical experiment is shown here in black.

tion we excluded sequences with high CG content from our work.

For each template, we measured the kinetics of binding of a series of complementary oligonucleotides of different lengths at concentrations typically ranging from ≈0.5 μM to ≈5 μM as shown in Figure 2A. The association rates determined through least squares fitting to a second order reaction scheme yielded the values shown in Figure 2B, where a positive dependence of the hybridization rates on oligonucleotide length can be observed.

To model RNA binding kinetics, we used two different approaches: an analytic treatment based on a simplified model and a more exact numerical treatment of the complete process of oligonucleotide hybridization; both approaches were developed based on previously published work in which annealing is considered to begin with the formation of a single base pair, followed by the subsequent formation (or loss) of additional base pairs (26,38,39,61).

In our simplified model, we treat the binding of two oligonucleotides as a multiple-step process: the oligonucleotides first diffuse and collide in the proper orientation to form an initial base pair with a bimolecular collision rate k_{bi}. After the initial contact is formed, new base pairs can be progressively added or removed, with the formation of new base pairs proceeding at a fixed unimolecular rate k_{uni}. Reverse steps for removal of a base pair proceed with a rate $k_{uni} \cdot e^{\Delta G_{j+1\to j}/RT}$, with ΔG being the free energy difference between the state with $j+1$ bound base pairs and the adjacent state with j bound base pairs. Analogously, the rate of detachment of two strands when connected by a single base pair is given by $k_{bi} \cdot e^{\Delta G_{1\to 0}/RT}$, which is equivalent to $k_{bi} \cdot K_D$, with K_D being the dissociation constant for the two strands linked only by the single base pair. This whole process can be expressed using the following reaction scheme:

$$A + B \underset{k_{bi}\cdot K_D}{\overset{k_{bi}}{\rightleftarrows}} AB_1 \underset{k_{uni}\cdot e^{\Delta G_{2\to 1}/RT}}{\overset{k_{uni}}{\rightleftarrows}} AB_2 \underset{k_{uni}\cdot e^{\Delta G_{3\to 2}/RT}}{\overset{k_{uni}}{\rightleftarrows}} \dots \underset{k_{uni}\cdot e^{\Delta G_{L\to L-1}/RT}}{\overset{k_{uni}}{\rightleftarrows}} AB_L \quad (5)$$

where AB_j represents a bound state between the A and B strands linked by j base pairs.

From this reaction scheme, under the assumption of a nucleation-limited process, we can define k_{on} as the following summation for every ith initial contact over the total number of initial contacts (N), with N equal to the length of the shortest of the two strands that attempt annealing (see Supplementary Data 7 and Supplementary Equations S1–S4 for complete derivation):

$$k_{on} \approx \sum_{i=1}^{N} k_{bi} \cdot \frac{2 \cdot k_{uni}}{2 \cdot k_{uni} + k_{bi} \cdot K_D^i} \quad (6)$$

Fitting our experimentally derived k_{on} data with this analytical model allowed us to reproduce measured k_{on} values with good accuracy (Pearson coefficient = 0.94) and no outliers, with the best fitting k_{bi} equal to 1.6×10^7 M^{-1} s^{-1} and k_{uni} equal to 2.1×10^5 s^{-1}. It is important to note that, while our function is sensitive to the value of k_{uni}, it is not equally constrained by k_{bi}, generating good predictions even when k_{bi} is much larger than our best fitting value. It follows that even if these parameters can be used to calculate hybridization rates, their physical interpretation should be treated with caution.

Since the computation of the binding energy and K_{D} for every possible initial contact is complicated by contributions from neighboring but non-paired nucleotides, we present here a simple function to approximate it starting from the CG content of the complementary stretch, that can be readily used without additional tools. To do so, we assume that a single average value describes the probability of successful annealing when any among the possible initial N contacts are formed, where N is the length of the shortest of the two complementary strands having CG content equal to f_{CG}. In this case, we can calculate a length-normalized k_{on}

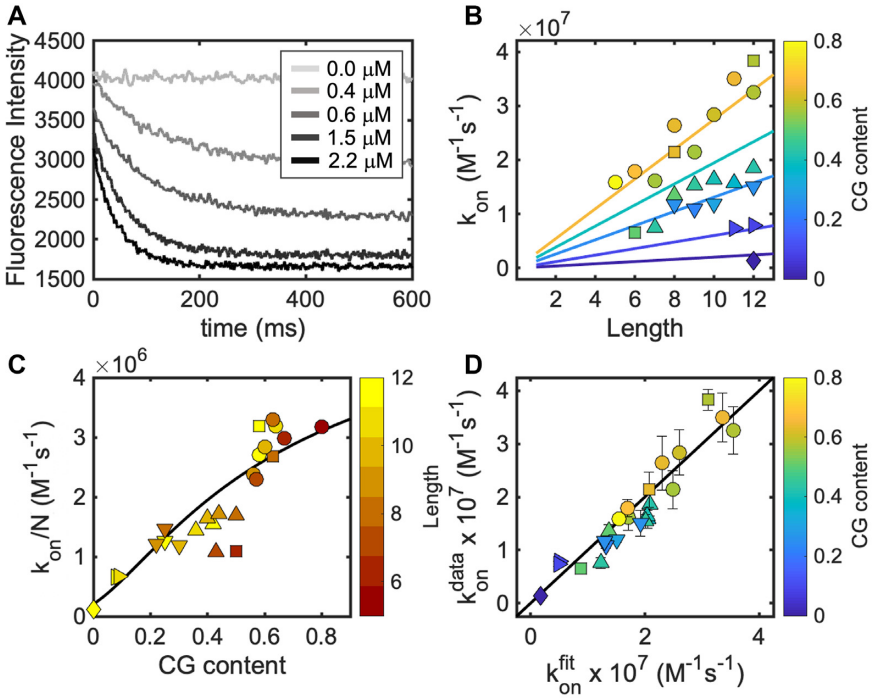

Figure 2. Association kinetics of short RNA oligonucleotides. (**A**) Example of hybridization of the 8 nt long oligonucleotide K8 to the KC* template sequence (at 0.5 μM) as a function of its concentration. (**B**) Experimentally derived k_{on} for short oligonucleotides show a linear dependence on their length. Different symbols are used for experiments performed using different templates as specified in Supplementary Table S2. Continuous lines represent predicted k_{on} values computed from Equation (7) with K_D calculated from Equation (8). Line colors match colormap on right. (**C**) length-normalized hybridization rates (M^{-1} s^{-1}) and approximating function computed form Equations (7) and (8). Deviations from the black line are likely due to sequence complexity not captured by an average K_D. (**D**) Correlation plot for calculated k_{on} using best fit k_{bi} and k_{uni} versus experimentally derived k_{on}. k_{on}^{fit} is computed using Equation (6) with $k_{bi} = 1.6 \times 10^7$ M^{-1} s^{-1} and $k_{uni} = 2.1 \times 10^5$ s^{-1}. Error bars are shared among data in panels B, C and D and only shown in the latter for clarity.

as follows:

$$\frac{k_{on(f_{CG})}}{N} \approx k_{bi} \cdot \frac{2 \cdot k_{uni}}{2 \cdot k_{uni} + k_{bi} \cdot K_{D(f_{CG})}} \quad (7)$$

Since there is no straightforward function to compute a priori the free energy of a base pair, we provide here an empirical function that relates the CG content of the RNA strand to the apparent ΔG of its initial contacts, extracted by expressing Equation (7) as a function of K_D and fitting our dataset. The following resulting equation can be used to compute K_D in Equation (7):

$$\Delta G\,(f_{CG}) = \frac{-2.27 \frac{kcal}{mol} \cdot f_{CG}}{0.24 + f_{CG}} \quad (8)$$

k_{on} parameters computed using Equations (7) and (8) are plotted as straight lines in Figure 2B, producing predictions that correlate well with experimentally derived k_{on} data (Pearson coefficient = 0.91).

With a simplified analytical model in hand, we tested our experimentally derived data against a more complete stochastic model that accounts for bulges and internal loops during the hybridization. Bulges and internal loops are expected to some degree according to coarse-grained simulations of repetitive sequences (32,40,62). We first tested the publicly available version of Multistrand (39), a software that models the annealing process as a trajectory over the complete Markov-chain of bound states among two strands, based upon a previously developed algorithm for RNA folding by Flamm *et al.* (63). Every transition between two states differs by a single nucleotide—either removed or added—and every state is characterized by a defined pattern of bound nucleotides, whose free energy is calculated using NUPACK, a powerful software package that computes the thermodynamics of systems containing multiple oligonucleotides (64). One advantage of such an approach is that it explicitly accounts for every step of the annealing process and should give a reasonable success rate in producing the aligned minimum free-energy structure when starting from an off-register initiation event. We found that even though the overall trend of our data is reproduced (Pearson coefficient = 0.66), there are many poorly predicted data points. In particular, the fCG = 0 sequence has a highly (20-fold) overestimated k_{on}. Given the low complexity of this sequence and the high number of off-register sites available to initiate its hybridization, the outcome of Multistrand suggests that the success rate for off-register trajectories may be overestimated with the current parametrization. We decided to re-parametrize the model, implementing our own version of the algorithm in Python.

One important difference between the reaction scheme employed in stochastic simulations on nucleic acids by Schaeffer and Flamm and the simplified analytical model

previously established in Equation (5) is the implementation of the Metropolis rate method for zippering. While in the simplified model every base pair addition will occur at a fixed forward rate, in the stochastic simulation the rate for any transition between two states that does not result in the detachment of the two strands is equal to k_{uni} for $\Delta G < 0$ (favorable transitions) or $k_{\text{uni}} * e^{(-\Delta G/RT)}$ for $\Delta G > 0$ (unfavorable transitions). To obtain hybridization rates through stochastic simulations, we implemented the 'first-step' approach established by Schaeffer (39): the two oligonucleotides were initialized in a complex connected by a single base pair picked randomly among the N possible initial in-register and off-register paired nucleotides, with N differing from the analytical treatment because it can be larger than the length of the shortest of the two hybridizing strands. The system then evolves using the Gillespie algorithm (65), picking a new state among all possible states, excluding secondary structures, that differ by a single base pair with a probability scaling on their rates. This process was iterated to produce a trajectory with the code stopping if the complex successfully reaches its minimum free-energy structure, or failing when the two strands detach.

By doing a grid search, we found the best fitting k_{bi} equal to 3.5×10^6 M^{-1} s^{-1} and k_{uni} equal to 4.25×10^5 s^{-1} (Pearson coefficient = 0.91), producing the simulated k_{on} values plotted against our experimental data in Figure 3A (see Supplementary Data 7 and Supplementary Equations S5–S6 for determination of hybridization rates).

Among the advantages of solving the complete Markov-chain for oligonucleotide annealing is that we could look in more detail at the role of in-register and off-register binding events (32,62) and their trajectories. Here we discuss annealing pathways considered as the set of unique binding configurations encountered during a trajectory, each one with a certain free-energy difference calculated from the reference unbound state and a minimal formation time, defined as the time elapsed from the initial contact until the given state is first reached. Looking at an in-register pathway (Figure 3B) we can see that it is a smooth function, with the annealing process resolving within a few microseconds. In contrast, with off-register contacts, we generally predicted a low success rate in hybridizing if the sequences have a low CG content (<0.5) or they are relatively short (<8nt). For longer sequences with high CG content (≥0.5), we have found a non-negligible probability of producing the minimum free-energy structure starting from an off-register site, as long as the offset is not too extreme. These trajectories can account for up to ∼60% of the total hybridization rate, with zippering passing through a more complex pathway with kinetic traps visible as in the case depicted in Figure 3C, where a pathway leading to the correct annealing of two strands starting from an off-register nucleation is plotted as an energy-time function.

In summary, our analysis of the annealing process of two short RNA oligonucleotides has allowed us to construct a simple model that returns accurate values of k_{on} using just two parameters (length and CG fraction), constituting the first step in producing a practical predictive model for the evolution over time of more complex systems of short RNA oligonucleotides.

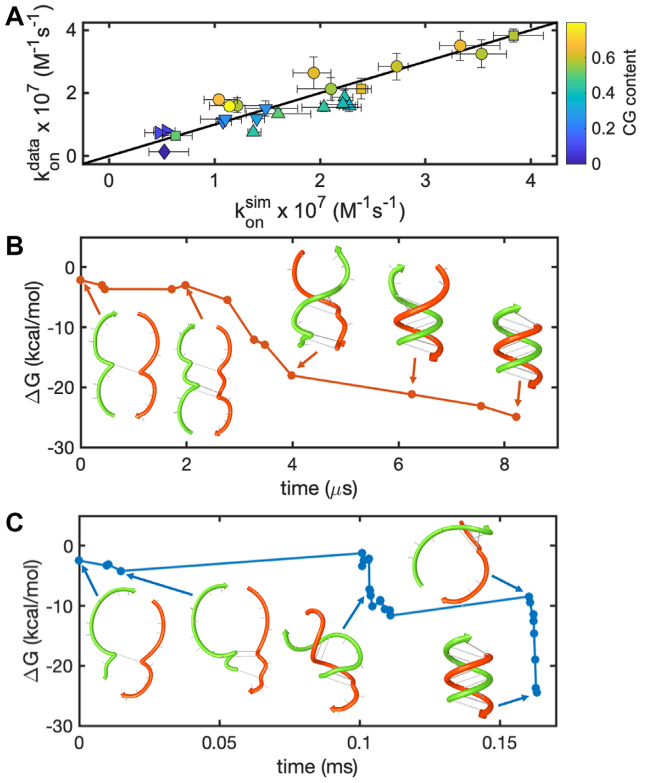

Figure 3. Stochastic simulations for the association kinetics of short RNA oligonucleotides. (**A**) Correlation plot for calculated k_{on} versus experimentally determined k_{on}. Calculated values were produced using $k_{\text{bi}} = 3.5 \times 10^6$ M^{-1} s^{-1} and $k_{\text{uni}} = 4.25 \times 10^5$ s^{-1}, sampling each initial binding site ten times and averaging over three replicates. Error bars on the x-axis represent standard deviations between the three replicates. Symbols used here refer to the same templates as the symbols used in Figure 2. (B, C) Annealing pathways for in-register (**B**) and off-register (**C**) contacts. In the first case we see a smooth pathway to the energy minimum, for the latter case the pathway shows kinetic traps solved through the correct realignment of the two strands. As soon as the two strands become correctly aligned, they zipper in a few microseconds. Selected pairing configurations along the two pathways are sketched using NUPACK web-based utilities to give a clearer representation of the different stages leading to a successful annealing.

Dissociation kinetics of short oligoribonucleotides

After the initial configuration of double helices in a mixture of short RNA oligonucleotides is established according to the hybridization rates for all pairs of strands, the duplexes will eventually separate after a certain amount of time governed by their k_{off}. This rate is crucial in determining the equilibration time-scale for any complex mixture and it is explored in this section.

Having measured k_{on} for our set of oligonucleotides and shown that these measured rates are in good agreement with a simple predictive model, we proceeded to determine the corresponding k_{off} rates from their dissociation constants as calculated from the NN model for oligonucleotides. This approach is reliable as long as duplex formation can be approximated as a two-states process, as already validated in the field of DNA hybridization (27,31). Before proceeding, we checked the accuracy of NN predictions of K_{D} for

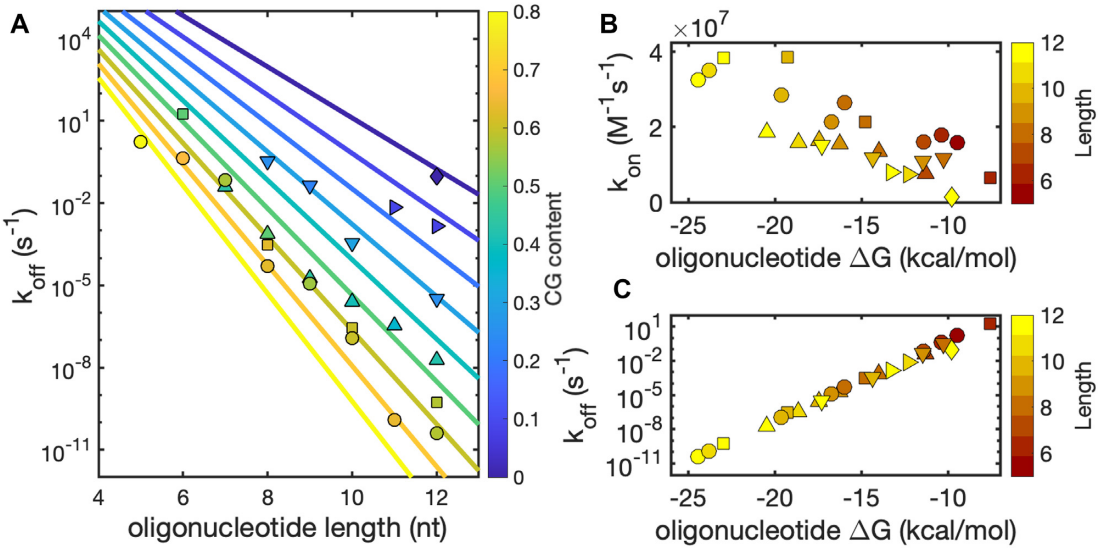

Figure 4. Kinetics of dissociation and association of short RNA oligonucleotides. k_{on} was measured as previously described in this work, while k_{off} was calculated as $k_{on} \cdot K_D$. (**A**) k_{off} for the oligonucleotides studied in this work as a function of length shows a poor correlation and emphasizes the dramatic scaling and role of CG content. Straight lines are predictions calculated from Equations (9) and (10). (**B**) k_{on} as a function of ΔG shows a linear dependence. (**C**) k_{off} as a function of ΔG shows a strong exponential correlation.

RNA duplexes at room temperature, since systematic discrepancies on the order of ∼1 kcal/mol are often found in studies on DNA (27,29,31). We obtained NN-predictions of the ensemble free energy (which includes the effects of frayed termini and other suboptimal configurations) by performing simulated melting experiments with NUPACK. We compared these values with our experimental data obtained by measuring the free energy of binding in a subset of our sequences through (i) direct K_D measurement through static binding experiments at room temperature *via* fluorescence spectroscopy, (ii) UV melting experiments, (iii) fluorescence melting experiments and (iv) kinetic binding measurements, finding that our measured ensemble free energy agrees very well with the predictions of the NN model (Pearson coefficient = 0.98, see Supplementary Figures S4 and S5).

Since NN-derived ΔGs are therefore reliable for the sequences used in this work, we calculated the k_{off} for all of our studied oligonucleotides as $k_{on} \cdot K_D$, where K_D was computed as $e^{\Delta G/(R \cdot T)}$ and ΔG is the free energy difference between the unbound and bound state. Data obtained in this way are presented as a function of oligonucleotide length (Figure 4A) and of the free energy of binding as a comparison with k_{on} (Figure 4B and C).

As expected, we observe a strong exponential correlation of k_{off} with ΔG, in agreement with the literature and the linear relationship between the activation energy for dissociation and ΔG (29,33). An equally clear linear correlation of k_{on} with ΔG can be observed, which is a direct consequence of the relationship stated in Equation (6). It follows that the main contributor to the overall duplex stability of the RNA oligonucleotide is k_{off}, whose values vary by roughly 10 orders of magnitude for the oligonucleotides studied in this work.

To better understand the relationship of k_{off} to length and CG content, we can express the activation energy barrier for dissociation using the Eyring equation (29,33,61):

$$k_{off} = \kappa \frac{k_B T}{h} e^{\frac{-\Delta G^{\ddagger}}{RT}} \quad (9)$$

where κ is the transmission coefficient (assumed to be one), k_B is Boltzmann's constant (1.38×10^{-23} J/K), T is the temperature in Kelvin, h is Planck's constant (6.63×10^{-34} J/Hz), R is the gas constant (1.99×10^{-3} kcal/mol/K) and $\Delta G^{\ddagger}$ is the activation free energy for the duplex dissociation.

We can define the activation energy for the dissociation of a duplex of length L as being the sum of the energy barriers to detach each base pair of the duplex, with a relative amplitude dependent on the identity of such pair, either A–U or C–G, each increasing the energy barrier by $\Delta G^{\ddagger}_{AU}$ and $\Delta G^{\ddagger}_{CG}$ respectively. An analogous approach has been recently employed by Rejali *et al.* using a NN model with 12 parameters fitted on a database of 43 sequences (29). In this work, we opt for a simpler parametrization, fitting our dataset using the Eyring equation with the activation energy in the following form:

$$\Delta G^{\ddagger} = L \cdot \left[f_{CG} \cdot \Delta G^{\ddagger}_{CG} + (1 - f_{CG}) \cdot \Delta G^{\ddagger}_{AU} \right] + \Delta G^{\ddagger}_{initiation} \quad (10)$$

This function offers an intuitive understanding of k_{off} and takes into account the linear dependence of k_{off} with length (on a log-linear scale) and the dependence of the slope on CG content, using only three free parameters. The best fit values for this model are 3.06 kcal/mol/bp for $\Delta G^{\ddagger}_{CG}$, 1.32 kcal/mol/bp for $\Delta G^{\ddagger}_{AU}$ and 3.18 kcal/mol for $\Delta G^{\ddagger}_{initiation}$. These parameters can be compared with the averaged values of Rejali *et al.* for DNA (29), equal to 2.24 kcal/mol for $\Delta G^{\ddagger}_{CG}$, 1.19 kcal/mol for $\Delta G^{\ddagger}_{AT}$ and 5.12 kcal/mol for $\Delta G^{\ddagger}_{initiation}$. Using this model, it is possible to extract ap-

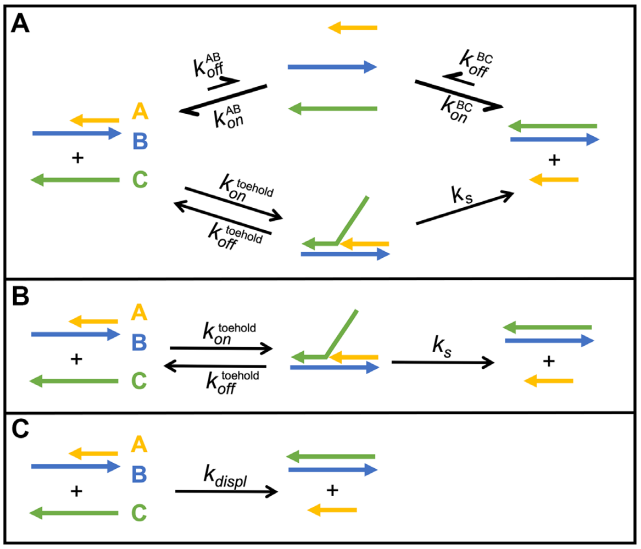

Figure 5. Reaction scheme for toehold-mediated strand displacement. (**A**) Complete reaction scheme. (**B**) Reaction scheme under the condition of negligible dissociation of duplexes. (**C**) Simplified reaction scheme valid when a rapid pre-equilibrium is established between the displacing oligonucleotide and the template.

proximate k_{off} values at any combination of length and CG content.

It is important to note that that the higher energy barrier for dissociation causes k_{off} in RNA to be much lower than that for DNA (33,43), with the parameter quickly reaching timescales much longer than common processes of interest in prebiotic and biological contexts. A consequence of this phenomenon is that equilibration processes for a solution of many oligonucleotides might take several years unless mechanisms to speed up this process intervene, such as strand displacement.

Toehold mediated strand displacement

Whenever a double helix and a single stranded oligonucleotide are present in solution, the equilibration will not be necessarily limited by the k_{off} of the duplex and could be greatly sped up due to strand displacement, a process that will inevitably happen in a competing scenario.

Strand displacement is a multi-step process where an oligonucleotide (A) bound to a template (B) is replaced by a third oligonucleotide (C). During the strand displacement process, the incoming displacing oligonucleotide binds to an available complementary stretch on the template, the so-called toehold, and undergoes branch migration, which leads to the release of the previously bound strand. We designed our sequences so that the displaced oligonucleotide is unlikely to rebind to the template since the latter is missing a useful toehold once bound to C. A complete reaction scheme for a system undergoing strand displacement should take into account association and dissociation of every species as depicted in Figure 5A.

If the strand displacement occurs on a timescale much shorter than the lifetime (τ) of the AB duplex, we can neglect the dissociation of both AB and BC ($\tau_{BC} > \tau_{AB}$ by design) so the reaction scheme can be reduced to a two-step reaction as shown in Figure 5B. Moreover, if the toehold binding equilibrates on a much faster time scale than the strand displacement reaction, which is reasonable for the short toehold lengths explored in this work, we can approximate the strand displacement process as a much simpler second order reaction, as shown in Figure 5C.

This simple model treats strand displacement as a bimolecular reaction with second order rate k_{displ} equal to $k_s \cdot K_A$, where K_A is the association constant of the toehold ($K_A = [ABC][AB]^{-1}[C]^{-1}$) and k_s the unimolecular rate of branch migration in the strand displacement process. In our experiments we monitored the evolution of the system by following the 2Ap fluorescence signal. Because [ABC] << [C] using short toeholds, even if 2Ap is placed in the toehold domain, the quenching measured in our experiment derives solely from formation of the BC duplex product. For the very same reason, because [ABC] is too low to be determined experimentally, we cannot directly measure K_A. To overcome this limitation we take advantage of NUPACK following the approach of Zhang & Winfree (48). Using NUPACK we can compute the free energy of user-defined structures and determine the difference in energy from states 'AB + C' to state 'ABC'. These values can be used to calculate the toehold binding energy ($\Delta G_{toehold}$) as $\Delta G_{AB+C} - \Delta G_{ABC}$ and the association constant for the toehold binding (K_A) equal to $e^{(\Delta G toehold/RT)}$.

For the case of strand displacement without an available toehold, we assume that strand displacement proceeds through the binding of C to the terminal nucleotides of the AB duplex, which are partially available due to the fraying of the duplex ends (66). In this case the strand displacement rate will be the sum of the rates for the displacements initiated at the two termini, $k_{displ}^{5'} = k_s \cdot K_A^{5'}$ and $k_{displ}^{3'} = k_s \cdot K_A^{3'}$, so that the overall phenomenon proceeds with an apparent rate $k_{displ} = k_s \cdot (K_A^{5'} + K_A^{3'})$, with an apparent binding energy of the invading strand to the toehold equal to:

$$\Delta G_{toehold}^{zero} = \ln\left(K_A^{5'} + K_A^{3'}\right) \cdot RT$$
$$= \ln\left(e^{\Delta G_{toehold}^{5'}/RT} + e^{\Delta G_{toehold}^{3'}/RT}\right) \cdot RT \qquad (11)$$

According to the model used here, a plot of k_{displ} versus $\Delta G_{toehold}$ should follow the exponential function $k_s \cdot e^{(-\Delta G toehold/RT)}$ that we used to fit our dataset obtained from a series of experiments performed using toeholds of lengths ranging from 0 nt to 5 nt and positioned either on 5′ or 3′ terminus of the template.

As the toehold binding strength increases, we find that the rate of strand displacement approaches a plateau as already observed in literature on DNA (48), while for toehold binding energies lower than 8 kcal/mol, our data yields a best fit of $k_s = 8.9\ s^{-1}$ (Figure 6B) with no significant difference between toeholds positioned on the 5′ or 3′ terminus of the template. The observed behavior is due to strand displacement entering a regime where a significant fraction of the toeholds is bound and the reaction becomes limited by k_s, so that the time traces that we measure are only apparent bimolecular reactions. That is, they will eventually equal the unimolecular rate k_s at high enough concentrations of the displacing oligonucleotide. We can estimate *a priori* what

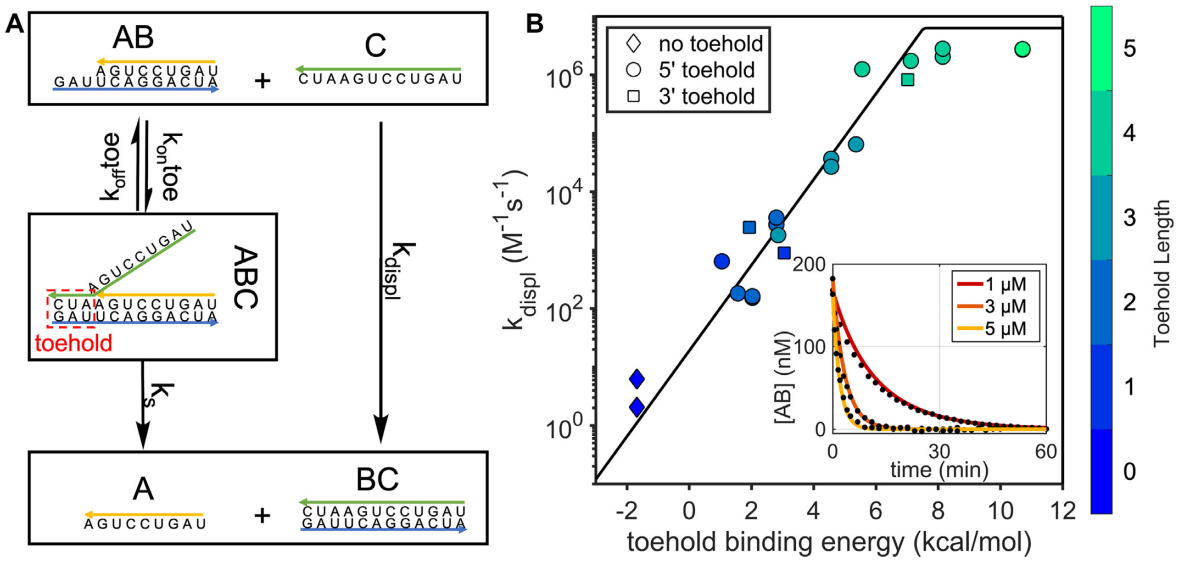

Figure 6. Strand displacement in short RNA oligonucleotides. (**A**) Reaction scheme for strand displacement. Under our conditions the two-step reaction can be simplified as a single bimolecular reaction with rate k_{displ}. (**B**) k_{displ} for short oligonucleotides has an exponential dependence on toehold binding affinity. Toehold binding affinity for the case with no toehold is derived as described in main text. The best fit for $k_s = 8.9\ s^{-1}$ is shown (black line). The saturation value has been set to $4.6 \times 10^6\ M^{-1}\ s^{-1}$, computed as described in the main text. Error bars are smaller than symbols size and are thus not shown. Inset depicts strand displacement time traces for different [C] for a 3nt long toehold ($\Delta G_{\mathrm{toehold}} = 2.86$ kcal/mol). Dots are experimental data and continuous lines are the result of a global fit for $k_{\mathrm{displ}} = 1.81 \times 10^3\ M^{-1}\ s^{-1}$.

the apparent bimolecular rate limit will be for strong toeholds at our working concentration by using a simple irreversible consecutive reaction model and calculating a predicted time trace reflecting the change in [A] concentration, with the first step being the toehold binding followed by the branch migration:

$$[A] = [AB] \cdot \left\{1 + \frac{1}{k_{on}[C]_0 - k_s} \cdot \left(k_s \cdot e^{-k_{on}[C]_0 \cdot t} - k_{on}[C]_0 \cdot e^{-k_s \cdot t}\right)\right\}, \quad (12)$$

where A is the displaced oligo, B the template, and C the displacing oligo, as previously defined.

Using $1 \times 10^7\ M^{-1}s^{-1}$ as a typical k_{on} value, k_s set to $8.9\ s^{-1}$ and $[C]_0$ to 1 μM, we obtain strand displacement time traces that resolve over a characteristic timescale of ≈0.2 s. This corresponds to an apparent bimolecular rate of $\approx 4.6 \times 10^6\ M^{-1}s^{-1}$ and reflects a saturating value under our working conditions. This value is comparable to the plateau of $\approx 3 \times 10^6\ M^{-1}\ s^{-1}$ that we find experimentally (see Figure 6B).

Approach to equilibrium of mixtures of oligonucleotides

Having measured the kinetic parameters governing the time-evolution of sets of oligonucleotides, we asked whether our approach to modeling the kinetics of annealing of short oligonucleotides could be used to predict the behavior of more complex mixtures. For each experiment we mixed four oligonucleotides: two short oligonucleotides (A_S, B_S) and two long oligonucleotides (A_L, B_L) designed so that all A-type and B-type molecules could bind to each other (see Figure 1 for sketch of the experimental design). We followed the competition between different binding interactions by monitoring the disappearance of the A_SB_L (and appearance of A_LB_L) duplex by placing 2Ap in a region of B_L where it is selectively quenched by A_L; the duplex of B_L with A_S leaves 2Ap in an overhang and therefore unpaired and unquenched. Once again, we exploited the properties of 2Ap to track the evolution over time of competing oligonucleotides. After mixing the single stranded oligonucleotides at a concentration on the order of 1 μM, the fluorescence signal rapidly drops due to the immediate formation of all possible duplexes: A_SB_S, A_LB_L, A_LB_S and A_SB_L coexisting in solution at different concentrations. This first process is too fast to be captured by hand mixing and thus is not shown in our time traces. After this initial rapid hybridization, the system is still out of equilibrium and evolves over time until it converges to its energy minimum, which corresponds to a state where all the duplexes in solution are A_SB_S and A_LB_L, while less stable duplexes A_LB_S and A_SB_L are absent. As the experiment progressed, we observed more and more quenching of the fluorescence emission of B_L due to its annealing to A_L. At all times after the initial rapid annealing, it must be true that $[A_SB_S] \approx [A_LB_L]$ and $[A_SB_L] \approx [A_LB_S]$. We could therefore determine the time course of each component of the system in the same experiment. In a typical measurement, we see the fluorescence dropping on a time scale that varies drastically depending on the oligonucleotide sequences and concentrations, with the presence of longer toeholds for strand displacement speeding up the equilibration in a way that can be tuned on time scales varying by orders of magnitude.

After the duplexes are quickly formed in solution with a relative abundance depending on their hybridization rate, different pathways are expected to drive the quenching of the fluorescent signal and equilibration of the system: (i) direct hybridization of A_L and B_L oligonucleotides as they become briefly available after detaching from their B_S and A_S counterparts (Figure 7A), a process in direct competition with (ii) strand displacement from available A_L in solution

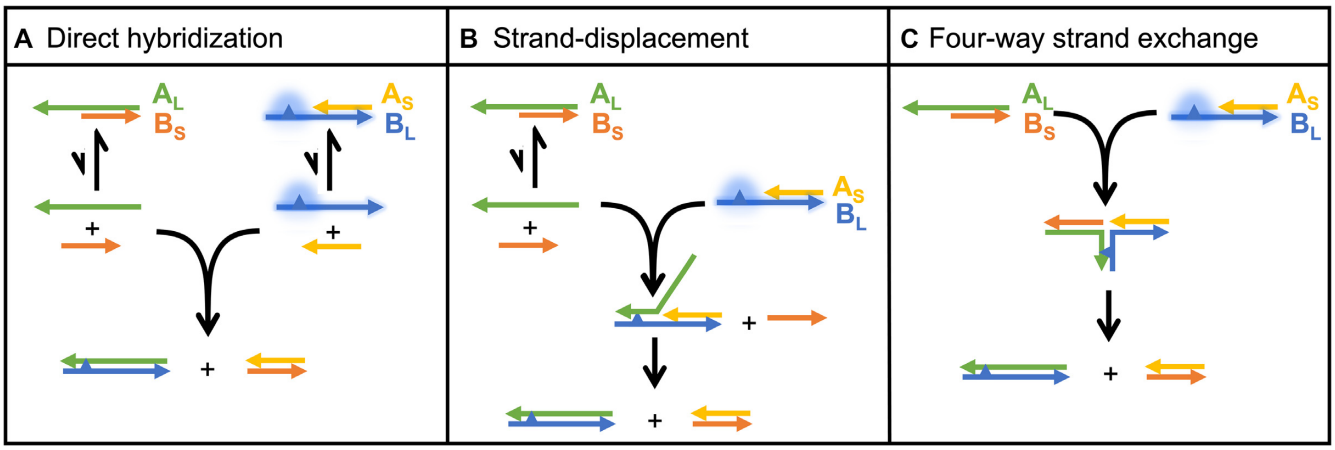

Figure 7. Equilibration pathways expected to determine the evolution of A_SB_L in complex mixtures of competing oligonucleotides.

due to an imbalance of the initial concentration of the different species or because they have become available after briefly detaching from their complementary B_S strand (Figure 7B) and (iii) four-way strand exchange, driving the direct substitution of two strands mediated by the interaction of the overhangs of two A_SB_L and A_LB_S duplexes, without the need for prior dissociation (Figure 7C).

According to the study by Dabby (67) on DNA four-way branch migration, we expected pathway (iii) to be extremely slow. To test whether this pathway was relevant in our case studies, we prepared mixtures of pre-annealed A_SB_L and A_LB_S at different total concentrations ($[A_SB_L] = [A_LB_S]$) and using sequences that would expose 4nt long complementary overhangs, the longest studied here. If a direct reaction exists between two duplexes with overhangs feeding pathway (iii), it will proceed through a two-step process similar to what has been previously established for strand displacement and described in the literature, namely, the formation of a 4-strand complex followed by strand exchange. The total rate of conversion from the initial to the final duplexes will depend on the fraction of oligos involved in the 4-strands complex, and this will have a concentration dependence that should be reflected in the reaction. However, our measurements clearly show a concentration-independence of the equilibration rates, allowing us to rule out the four-way branch exchange pathway (see Supplementary Figure S6). Our data set an upper limit for the rate of the four-way strand exchange process, which must be slower than $\approx 10^2$ M^{-1} s^{-1} for duplexes with 4nt long overhangs.

Having ruled out the third equilibration pathway, we focused on the first two, that should be well described by the parameters characterized in this work. To model our time course experiments, we first calculated k_{on}, k_{displ} and k_{off} as described in the previous sections, and then numerically solved a set of differential equations describing the system using the calculated parameters, varying them within their experimental error range (see Supplementary Data 9 for the complete system of differential equations). In Figure 8, we show experimental results compared with our predictions depicted as shaded areas. The time courses numerically computed using parameters calculated as described in this work do approximate the evolution of the different species, capturing equilibration times ranging from a few minutes up to several hours. For each case study, we determined the relative contribution of pathways (i) and (ii) toward equilibration after the first quick hybridization has been established, computed as the relative amount of A_LB_L produced through the two pathways and expressed as percentages in the right panel of Figure 8. We can describe two general cases for the equilibration of these mixtures: (i) for unbalanced mixtures having concentration of A_L larger than A_S, B_S and B_L (Figure 8B, C) we can in principle push the equilibration rate up to the branch migration rate (k_s) by increasing the toehold length and A_L concentration; for balanced mixtures (Figure 8A, D) the equilibration timescale shows a lower limit corresponding to the lifetime of A_S- and B_S-containing duplexes. The equilibration in such cases effectively coincides with the lifetime of the A_SB_L duplex ($\approx$1 h) under the conditions of a fast strand displacement (Figure 8A). This happens because for A_L and B_L strands to hybridize and drive the equilibration, they must first detach from their complementary B_S and A_S strand respectively, so that even with strand displacement occurring at its maximum rate equal to k_s, the equilibration will be limited by the dissociation of the short strands.

DISCUSSION

We have examined the kinetics of strand association and dissociation for RNA oligonucleotides with no significant secondary structures having CG content ranging from 0.0 to 0.8 and length between 5nt and 12nt. We have provided a simplified analytical model to predict RNA oligonucleotide binding kinetics and validated a Markov-chain model for the description of complete binding trajectories. Our analysis shows that RNA annealing is generally limited by the formation of just two in-register base pairs in the context of a nucleation-zipper model. Following this result, two base pairs constitute a nucleation site with a high probability of progressing to complete annealing *via* a fast zippering process. For relatively long oligonucleotides with a high CG content, stochastic simulations show that off-register trajectories become available, with a zippering timescale dominated by the realignment of the two strands. We find that

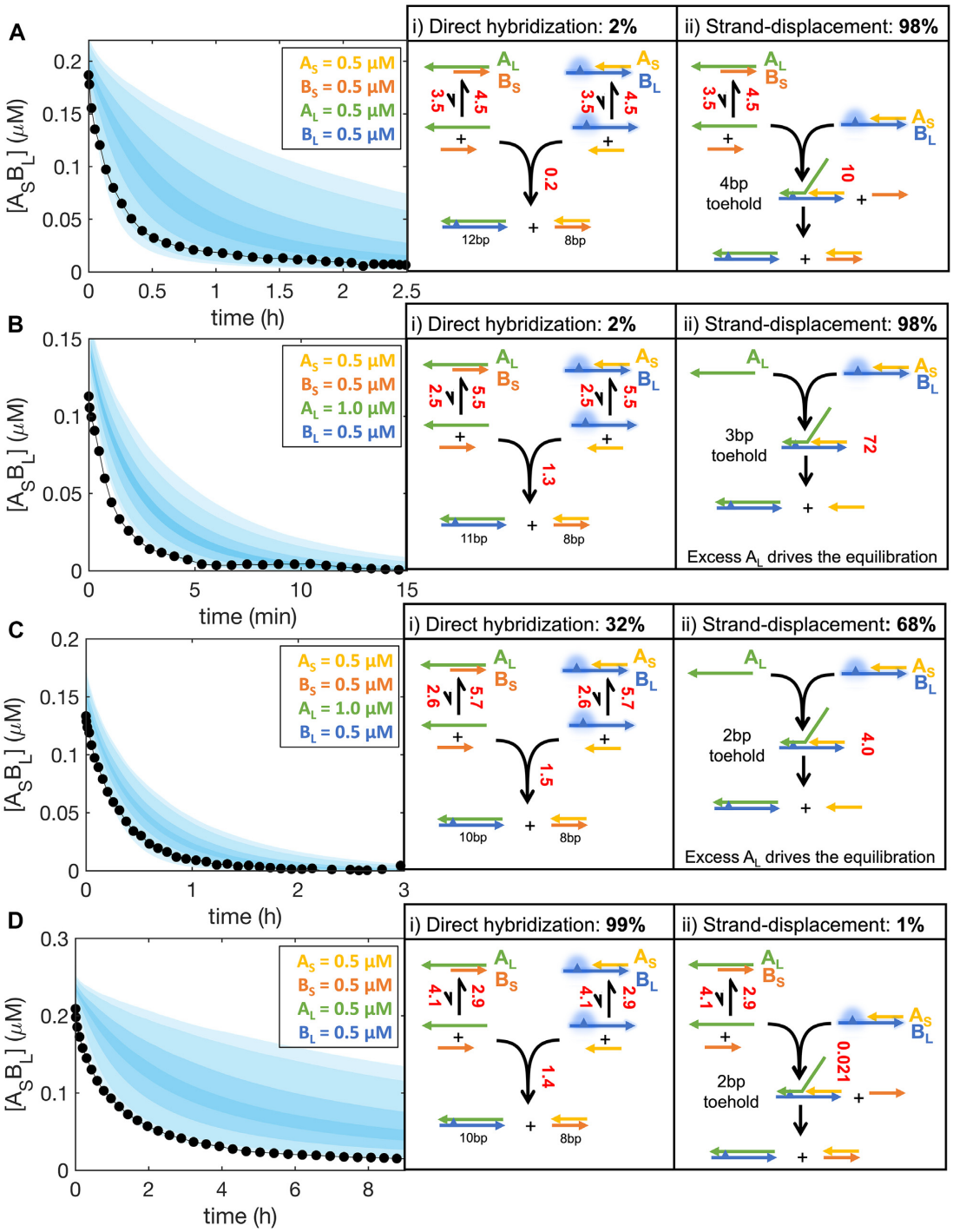

Figure 8. Examples of time traces for the evolution of A_SB_L in complex mixtures of competing oligonucleotides (black) compared with predictions from our modeling (shaded area). The measured signal is coming from the species A_SB_L and disappears when the solution is composed only by A_SB_S and A_LB_L following the pathways depicted on the right. For each pathway the relative contribution to equilibration is given as a percentage of A_LB_L produced after the initial rapid hybridization. Concentrations of A_S, A_L, B_S and B_L sequences used in each experiment are specified in every panel, with the color code for each strand matching the corresponding sketches on the right. Numbers in red next to each reaction arrow specify the corresponding instantaneous rates in units of nM/min, calculated right after the initial duplexes are established through the rapid hybridization of the free single stranded RNA (typically a few seconds after mixing). Shaded areas enclose trajectories calculated by resampling the kinetic parameters within their estimated confidence intervals and are drawn to envelope density probabilities equal to 95%, 90%, 60% and 20% of overlapping the real trajectory. Sequences used are specified in Supplementary Data 9.

our model is consistent with a single collision rate, supporting the hypothesis of Wetmur and Davidson (38) that any specific nucleation event will occur at a rate that is independent of the length of the oligonucleotide for short oligonucleotides (<100nt). For relatively short RNA strands the hybridization rate should increase with length because the number of in-register nucleation sites is proportional to length; thus k_{on} should grow at least in a linear fashion with oligonucleotide length. This dependence may even be super-linear if off-register nucleation sites become available, as discussed in our study of hybridization mechanisms. However, as lengths increase further (Wetmur and Davidson predict the threshold to be at ∼100 bp) the annealing rate is expected to grow only as the square root of the length due to additional interactions such as excluded volume effects that modify the accessibility of the contact sites. The increasing likelihood of secondary structures forming in longer single strands may also limit the annealing rate, although such effects are sequence dependent and difficult to predict. For this reason, we have dealt exclusively with oligonucleotides short enough that their secondary structure could be minimized by design.

Interestingly, although a sharp drop in hybridization rate for oligonucleotides with fewer than seven consecutive complementary nucleotides has been observed for both RNA and DNA, leading to the formulation of the so-called 'rule of seven' (1), we do not observe a discontinuity in on-rates vs. length at a length of 7, suggesting that a more complicated mismatch-dependent phenomenon might be at play in the work by Cisse and colleagues.

Our model is in agreement with the results of Rejali *et al.* (29) who showed that a NN-based prediction of k_{on} is unsuccessful. With the aim of providing a predictive algorithm for the hybridization kinetics of short RNA oligonucleotides, we have shown that it is possible to calculate reliable k_{on} values using Equations (7) and (8) as a function of length (N) and CG content, two parameters that are easily determined. With these k_{on} parameters in hand, we could compute the dissociation rate or k_{off} of short duplexes from the free energy of binding, calculated using the nearest-neighbor model for RNA. This step is critically dependent on the accuracy of the calculated ΔG, and we found that NUPACK predicted the experimentally measured values with reasonable accuracy. We find a clear dependence of k_{off} on oligonucleotide length and ΔG in agreement with existing literature (27,29,31,43), and we could model this dependence by relating the dissociation to the activation energy contribution coming from each A–U and C–G pair, in a similar fashion to that of Rejali *et al.* (29) This approach provides a way to quickly estimate dissociation kinetics for arbitrary sequences of short RNA oligonucleotides with negligible secondary structure, bypassing the tedious calculation of free energies of binding. As expected k_{off} for progressively longer RNA duplexes quickly reaches extremely long timescales, making equilibration of complex mixtures unlikely without thermal annealing or some other non-equilibrium relaxation phenomena.

To begin to explore the equilibration time-scales of more complex mixtures of oligoribonucleotides, we first explored strand displacement in RNA duplexes. Our experimental results are in agreement with the coarse-grained simulations of Šulc *et al.* (49) As previously shown for DNA, we also find that this process is controlled by the binding energy of the invading strand for the template toehold and we have determined a strand displacement rate k_{displ}, equal to $k_s \cdot K_A$, with a best fit of $k_s = 8.9\ s^{-1}$ for toeholds having $\Delta G < 7$ kcal/mol. Values of the same parameter measured for DNA range from $0.7\ s^{-1}$ to $7\ s^{-1}$ (48). Compared to strand displacement in DNA, where the rate plateaus at values close to k_{on} for a toehold length of ≈6 nt, the RNA strand displacement rate in our study approaches limiting values of $\approx 3 \times 10^6\ M^{-1}s^{-1}$ using a toehold of 4 nt (when the CG content of the toehold is ≈0.5) due to its stronger base pairing. This observation implies that equilibration processes in complex RNA systems will be generally faster than analogous DNA systems if they are dominated by strand displacement. For a more detailed analysis of strand displacement, branch migration can be modeled as a process with a unimolecular rate k_s that can be decomposed using the Gambler's Ruin analysis as equal to $k_{mig} \cdot N^{-2}$ (68), where k_{mig} is the first-order rate for one elementary branch migration step and N is the length of the duplex region being displaced by branch migration. According to Zhang *et al.* (48) k_{mig} in DNA at 1 M NaCl is $\approx 400\ s^{-1}$. For our data, we obtain k_{mig} equal to $\approx 565\ s^{-1}$, confirming that branch migration occurs at a similar rate for RNA.

Altogether, in this work we have provided a framework for the prediction of the hybridization kinetics of short oligoribonucleotides, and we have addressed the challenge of predicting the time-course evolution of mixtures of competing oligoribonucleotides. We have shown that our approach can be used to compute the evolution of equilibrating solutions containing multiple RNA sequences, which is ultimately dependent on two competing pathways of direct hybridization and strand displacement. While we could rule out the role of four-way strand exchange in our case studies, we can speculate here whether this pathway is going to be relevant when dealing with longer overhangs, that would result in an increase of the four-way strand exchange rate. If we keep the 12nt-long template fixed, we must reduce the duplexed stretch to elongate the overhang, and this will cause a drastic speed up of its k_{off} and reduction of its lifetime. Using as a reference the branch migration rates published by Nadine (67) and assuming the ones for RNA are comparable (as expected following our results for RNA strand displacement), it is reasonable to expect the equilibration of two duplexes with a long overhang to be ultimately dominated by the first two pathways previously described, so that four-way branch migration is effectively always negligible in mixtures of short RNA oligonucleotides.

By solving a simple system of differential equations, we could predict the lifetime and concentration of every component of the complex mixture, finding that the time required to anneal two complementary strands can be greatly extended by increasing the complexity of the system. While two oligonucleotides typically anneal in a fraction of a second, by varying the composition of a mixture of only four oligonucleotides we could easily manipulate the equilibration time from minutes up to several hours by creating kinetic traps. One important implication of this result is that complexes that are not expected to exist at equilibrium will

be present and long-lived when the system initializes in an out-of-equilibrium state. Surprisingly, we find that such duplexes can exist for timescales much longer than their k_{off}-determined lifetimes when their equilibration process is dominated by the direct hybridization pathway. Following this observation, we predict that a system composed by a large collection of duplexes exposing only short toeholds will undergo an exceptionally long equilibration process. This finding is particularly relevant in the context of models for the nonenzymatic replication of genetic information, which is believed to be driven by chemical reactions occurring in highly heterogeneous and complex mixtures of oligoribonucleotides (12,69).

Although slowing down equilibration processes is important for systems in which metastable duplexes are substrates of interest, a more desirable scenario is one in which a mixture quickly reaches a target equilibrium state, spending as little time as possible in unwanted kinetic traps that could lead to off-target reaction products. Our study sheds light on the best strategies to speed up equilibration when annealing is not a viable option, for example when dealing with temperature-sensitive systems containing unstable species. Generally, a balanced system of the type studied here can be predicted to equilibrate in less than a minute if the metastable duplex has a lifetime no longer than ∼1 s. Given the slow dissociation rate of RNA, even relatively short RNA duplexes will have a longer lifetime, so that it is critical for these systems to be designed in such a way that any metastable duplexes bear long toeholds, and that the mixtures are prepared with an excess of displacing oligonucleotides to quickly drive the conversion of the metastable duplexes to the desired configuration

Our work provides the basis for future studies on larger systems of competing strands, and to the design of systems with tunable equilibration rates for desired species. We suggest that an extension of the stochastic simulation for hybridization to account for the effect of coaxial stacking and more complex configurations on the kinetics of annealing, dissociation and strand displacement (70) will allow this approach to be extended to the study of complexes composed of many strands in large interacting networks of oligonucleotides.

DATA AVAILABILITY

Oligonucleotide sequences and mixtures studied here are listed in the online supplemental data file accompanying this article. Raw data are available at OSF.io: https://osf.io/b6kfy/?view_only=44ae68ca639b4566ae01f3cdb96b1886.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank Dian Ding, Shriyaa Mittal and Lijun Zhou for valuable discussions and for sharing their virtual circular genome sequence design. We thank Harry Aitken, Brennan Ashwood, Marco Buscaglia, Ben Colville, Victor Lelyveld, Aleksandar Radakovic, Kyle Strom, Andrei Tokmakoff and Longfei Wu for comments on the manuscript. M.T. would like to thank Giuliano Zanchetta for introducing him to nucleic acid kinetics and Daniel Duzdevich for his help in improving the manuscript.

FUNDING

National Science Foundation [CHE-2104708 to J.W.S.]; Simons Foundation [290363 to J.W.S.]; J.W.S. is an investigator of the Howard Hughes Medical Institute. Funding for open access charge: Simons Foundation.
Conflict of interest statement. None declared.

REFERENCES

1. Cisse,II, Kim,H. and Ha,T. (2012) A rule of seven in Watson-Crick base-pairing of mismatched sequences. *Nat. Struct. Mol. Biol.*, **19**, 623–627.
2. Ui-Tei,K., Nishi,K., Takahashi,T. and Nagasawa,T. (2012) Thermodynamic control of small RNA-mediated gene silencing. *Front. Genet.*, **3**, 101.
3. Burghgraeve,N., Simon,S., Barral,S., Fobis-Loisy,I., Holl,A.-C., Ponitzki,C., Schmitt,E., Vekemans,X. and Castric,V. (2020) Base-pairing requirements for small RNA-Mediated gene silencing of recessive self-incompatibility alleles in arabidopsis halleri. *Genetics*, **215**, 653–664.
4. Modi,S., Swetha,M.G., Goswami,D., Gupta,G.D., Mayor,S. and Krishnan,Y. (2009) A DNA nanomachine that maps spatial and temporal pH changes inside living cells. *Nat. Nanotechnol.*, **4**, 325–330.
5. Koos,B., Cane,G., Grannas,K., Lof,L., Arngarden,L., Heldin,J., Clausson,C.M., Klaesson,A., Hirvonen,M.K., de Oliveira,F.M. *et al.* (2015) Proximity-dependent initiation of hybridization chain reaction. *Nat. Commun.*, **6**, 7294.
6. Adleman,L.M. (1994) Molecular computation of solutions to combinatorial problems. *Science*, **266**, 1021–1024.
7. Parolini,L., Kotar,J., Di Michele,L. and Mognetti,B.M. (2016) Controlling self-assembly kinetics of DNA-functionalized liposomes using toehold exchange mechanism. *ACS Nano*, **10**, 2392–2398.
8. Dey,S., Fan,C., Gothelf,K.V., Li,J., Lin,C., Liu,L., Liu,N., Nijenhuis,M.A.D., Saccà,B., Simmel,F.C. *et al.* (2021) DNA origami. *Nat. Rev. Methods Primers*, **1**, 13.
9. Smith,G.P., Fraccia,T.P., Todisco,M., Zanchetta,G., Zhu,C., Hayden,E., Bellini,T. and Clark,N.A. (2018) Backbone-free duplex-stacked monomer nucleic acids exhibiting Watson-Crick selectivity. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, E7658–E7664.
10. Kudella,P.W., Tkachenko,A.V., Salditt,A., Maslov,S. and Braun,D. (2021) Structured sequences emerge from random pool when replicated by templated ligation. *P. Natl. Acad. Sci. U.S.A.*, **118**, e2018830118.
11. Rosenberger,J.H., Goppel,T., Kudella,P.W., Braun,D., Gerland,U. and Altaner,B. (2021) Self-assembly of informational polymers by templated ligation. *Phys. Rev. X*, **11**, 031055.
12. Zhou,L., Ding,D. and Szostak,J.W. (2021) The virtual circular genome model for primordial RNA replication. *RNA*, **27**, https://rnajournal.cshlp.org/content/early/2020/10/07/rna.077693.120.
13. Zhou,L., O'Flaherty,D.K. and Szostak,J.W. (2020) Assembly of a ribozyme ligase from short oligomers by nonenzymatic ligation. *J. Am. Chem. Soc.*, **142**, 15961–15965.
14. SantaLucia,J. and Hicks,D. (2004) The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, **33**, 415–440.
15. Xia,T., SantaLucia,J., Burkard,M.E., Kierzek,R., Schroeder,S.J., Jiao,X., Cox,C. and Turner,D.H. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry*, **37**, 14719–14735.
16. Genot,A.J., Bath,J. and Turberfield,A.J. (2011) Reversible logic circuits made of DNA. *J. Am. Chem. Soc.*, **133**, 20080–20083.
17. Bellini,T., Zanchetta,G., Fraccia,T.P., Cerbino,R., Tsai,E., Smith,G.P., Moran,M.J., Walba,D.M. and Clark,N.A. (2012) Liquid crystal self-assembly of random-sequence DNA oligomers. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 1110–1115.

18. Zhou,L., O'Flaherty,D.K. and Szostak,J.W. (2020) Template-Directed copying of RNA by Non-enzymatic ligation. *Angew. Chem. Int. Ed Engl.*, **59**, 15682–15687.
19. Walton,T., Zhang,W., Li,L., Tam,C.P. and Szostak,J.W. (2019) The mechanism of nonenzymatic template copying with imidazole-activated nucleotides. *Angew. Chem. Int. Ed Engl.*, **58**, 10812–10819.
20. Wu,L.F., Su,M., Liu,Z., Bjork,S.J. and Sutherland,J.D. (2021) Interstrand aminoacyl transfer in a tRNA acceptor stem-overhang mimic. *J. Am. Chem. Soc.*, **143**, 11836–11842.
21. Gao,Y., Wolf,L.K. and Georgiadis,R.M. (2006) Secondary structure effects on DNA hybridization kinetics: a solution versus surface comparison. *Nucleic Acids Res.*, **34**, 3370–3377.
22. Schreck,J.S., Ouldridge,T.E., Romano,F., Šulc,P., Shaw,L.P., Louis,A.A. and Doye,J.P.K. (2015) DNA hairpins destabilize duplexes primarily by promoting melting rather than by inhibiting hybridization. *Nucleic Acids Res.*, **43**, 6181–6190.
23. Ouldridge,T.E., Sulc,P., Romano,F., Doye,J.P. and Louis,A.A. (2013) DNA hybridization kinetics: zippering, internal displacement and sequence dependence. *Nucleic. Acids. Res.*, **41**, 8886–8895.
24. Craig,M.E., Crothers,D.M. and Doty,P. (1971) Relaxation kinetics of dimer formation by self complementary oligonucleotides. *J. Mol. Biol.*, **62**, 383–401.
25. Porschke,D. and Eigen,M. (1971) Co-operative non-enzymic base recognition. 3. Kinetics of the helix-coil transition of the oligoribouridylic–oligoriboadenylic acid system and of oligoriboadenylic acid alone at acidic pH. *J. Mol. Biol.*, **62**, 361–381.
26. Wako,H. and Saitô,N. (1978) Statistical mechanical theory of the protein conformation. I. General considerations and the application to homopolymers. *J. Phys. Soc. Jpn.*, **44**, 1931–1938.
27. Dupuis,N.F., Holmstrom,E.D. and Nesbitt,D.J. (2013) Single-molecule kinetics reveal cation-promoted DNA duplex formation through ordering of single-stranded helices. *Biophys. J.*, **105**, 756–766.
28. Bonnet,G., Krichevsky,O. and Libchaber,A. (1998) Kinetics of conformational fluctuations in DNA hairpin-loops. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 8602–8606.
29. Rejali,N.A., Ye,F.D., Zuiter,A.M., Keller,C.C. and Wittwer,C.T. (2021) Nearest-neighbour transition-state analysis for nucleic acid kinetics. *Nucleic Acids Res.*, **49**, 4574–4585.
30. Shen,Y.Q., Kuznetsov,S.V. and Ansari,A. (2001) Loop dependence of the dynamics of DNA hairpins. *J. Phys. Chem. B*, **105**, 12202–12211.
31. Menssen,R.J. and Tokmakoff,A. (2019) Length-dependent melting kinetics of short DNA oligonucleotides using temperature-jump IR spectroscopy. *J. Phys. Chem. B*, **123**, 756–767.
32. Jones,M.S., Ashwood,B., Tokmakoff,A. and Ferguson,A.L. (2021) Determining sequence-dependent DNA oligonucleotide hybridization and dehybridization mechanisms using coarse-grained molecular simulation, Markov state models, and infrared spectroscopy. *J. Am. Chem. Soc.*, **143**, 17395–17411.
33. Rauzan,B., McMichael,E., Cave,R., Sevcik,L.R., Ostrosky,K., Whitman,E., Stegemann,R., Sinclair,A.L., Serra,M.J. and Deckert,A.A. (2013) Kinetics and thermodynamics of DNA, RNA, and hybrid duplex formation. *Biochemistry*, **52**, 765–772.
34. Yuan,B.F., Zhuang,X.Y., Hao,Y.H. and Tan,Z. (2008) Kinetics of base stacking-aided DNA hybridization. *Chem. Commun.*, 6600–6602.
35. Nava,G., Ceccarello,E., Giavazzi,F., Salina,M., Damin,F., Chiari,M., Buscaglia,M., Bellini,T. and Zanchetta,G. (2016) Label-free detection of DNA single-base mismatches using a simple reflectance-based optical technique. *Phys. Chem. Chem. Phys.*, **18**, 13395–13402.
36. Hertel,S., Spinney,R.E., Xu,S.Y., Ouldridge,T.E., Morris,R.G. and Lee,L.K. (2022) The stability and number of nucleating interactions determine DNA hybridization rates in the absence of secondary structure. *Nucleic Acids Res.*, **50**, 7829–7841.
37. Schickinger,M., Zacharias,M. and Dietz,H. (2018) Tethered multifluorophore motion reveals equilibrium transition kinetics of single DNA double helices. *Proc. Natl. Acad. Sci.*, **115**, E7512–E7521.
38. Wetmur,J.G. and Davidson,N. (1968) Kinetics of renaturation of DNA. *J. Mol. Biol.*, **31**, 349–370.
39. Schaeffer,J.M. (2013) In: *Dissertation (Ph.D.)*. California Institute of Technology.
40. Hinckley,D.M., Lequieu,J.P. and de Pablo,J.J. (2014) Coarse-grained modeling of DNA oligomer hybridization: length, sequence, and salt effects. *J. Chem. Phys.*, **141**, 035102.
41. Zhang,J.X., Fang,J.Z., Duan,W., Wu,L.R., Zhang,A.W., Dalchau,N., Yordanov,B., Petersen,R., Phillips,A. and Zhang,D.Y. (2018) Predicting DNA hybridization kinetics from sequence. *Nat. Chem.*, **10**, 91–98.
42. Sikorav,J.L., Orland,H. and Braslau,A. (2009) Mechanism of thermal renaturation and hybridization of nucleic acids: kramers' process and universality in watson-crick base pairing. *J. Phys. Chem. B*, **113**, 3715–3725.
43. Woodside,M.T., Behnke-Parks,W.M., Larizadeh,K., Travers,K., Herschlag,D. and Block,S.M. (2006) Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 6190–6195.
44. Choi,H.M., Beck,V.A. and Pierce,N.A. (2014) Next-generation in situ hybridization chain reaction: higher gain, lower cost, greater durability. *ACS Nano*, **8**, 4284–4294.
45. Dirks,R.M. and Pierce,N.A. (2004) Triggered amplification by hybridization chain reaction. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 15275–15278.
46. Zhao,S., Yu,L., Yang,S., Tang,X., Chang,K. and Chen,M. (2021) Boolean logic gate based on DNA strand displacement for biosensing: current and emerging strategies. *Nanoscale Horiz*, **6**, 298–310.
47. Zhou,L., Kim,S.C., Ho,K.H., O'Flaherty,D.K., Giurgiu,C., Wright,T.H. and Szostak,J.W. (2019) Non-enzymatic primer extension with strand displacement. *Elife*, **8**, e51888.
48. Zhang,D.Y. and Winfree,E. (2009) Control of DNA strand displacement kinetics using toehold exchange. *J. Am. Chem. Soc.*, **131**, 17303–17314.
49. Hong,F. and Sulc,P. (2019) An emergent understanding of strand displacement in RNA biology. *J. Struct. Biol.*, **207**, 241–249.
50. Liu,H., Hong,F., Smith,F., Goertz,J., Ouldridge,T., Stevens,M.M., Yan,H. and Šulc,P. (2021) Kinetics of RNA and RNA:DNA hybrid strand displacement. *ACS Synth. Biol.*, **10**, 3066-3073.
51. Tataurov,A.V., You,Y. and Owczarzy,R. (2008) Predicting ultraviolet spectrum of single stranded and double stranded deoxyribonucleic acids. *Biophys. Chem.*, **133**, 66–70.
52. Xu,D.G. and Nordlund,T.M. (2000) Sequence dependence of energy transfer in DNA oligonucleotides. *Biophys. J.*, **78**, 1042–1058.
53. Jahnke,K., Grubmuller,H., Igaev,M. and Gopfrich,K. (2021) Choice of fluorophore affects dynamic DNA nanostructures. *Nucleic Acids Res.*, **49**, 4186–4195.
54. Owczarzy,R., Moreira,B.G. and You,Y. (2004) Effects of fluorescent dyes, quenchers, and dangling ends on the thermodynamic stability of DNA duplexes. *Biophys. J.*, **86**, 310a.
55. You,Y., Tataurov,A.V. and Owczarzy,R. (2011) Measuring thermodynamic details of DNA hybridization using fluorescence. *Biopolymers*, **95**, 472–486.
56. Narayanan,R., Zhu,L., Velmurugu,Y., Roca,J., Kuznetsov,S.V., Prehna,G., Lapidus,L.J. and Ansari,A. (2012) Exploring the energy landscape of nucleic acid hairpins using laser temperature-jump and microfluidic mixing. *J. Am. Chem. Soc.*, **134**, 18952–18963.
57. Law,S.M., Eritja,R., Goodman,M.F. and Breslauer,K.J. (1996) Spectroscopic and calorimetric characterizations of DNA duplexes containing 2-aminopurine. *Biochemistry*, **35**, 12329–12337.
58. Hall,K.B. and Williams,D.J. (2004) Dynamics of the IRE RNA hairpin loop probed by 2-aminopurine fluorescence and stochastic dynamics simulations. *RNA*, **10**, 34–47.
59. Gellert,M., Lipsett,M.N. and Davies,D.R. (1962) Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. U.S.A.*, **48**, 2013–2018.
60. Ralph,R.K., Connors,W.J. and Khorana,H.G. (1962) Secondary structure and aggregation in deoxyguanosine oligonucleotides. *J. Am. Chem. Soc.*, **84**, 2265–2266.
61. Menssen,R.J., Kimmel,G.J. and Tokmakoff,A. (2021) Investigation into the mechanism and dynamics of DNA association and dissociation utilizing kinetic monte carlo simulations. *J. Chem. Phys.*, **154**, 045101.
62. Xiao,S., Sharpe,D.J., Chakraborty,D. and Wales,D.J. (2019) Energy landscapes and hybridization pathways for DNA hexamer duplexes. *J. Phys. Chem. Lett.*, **10**, 6771–6779.

63. Flamm,C., Fontana,W., Hofacker,I.L. and Schuster,P. (2000) RNA folding at elementary step resolution. *RNA (New York, N.Y.)*, **6**, 325–338.
64. Zadeh,J.N., Steenberg,C.D., Bois,J.S., Wolfe,B.R., Pierce,M.B., Khan,A.R., Dirks,R.M. and Pierce,N.A. (2011) NUPACK: analysis and design of nucleic acid systems. *J. Comput. Chem.*, **32**, 170–173.
65. Gillespie,D.T. (1977) Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, **81**, 2340–2361.
66. Sanstead,P.J., Stevenson,P. and Tokmakoff,A. (2016) Sequence-dependent mechanism of DNA oligonucleotide dehybridization resolved through infrared spectroscopy. *J. Am. Chem. Soc.*, **138**, 11792–11801.
67. Nadine,D. (2013) In: *Dissertation (Ph.D.)*. California Institue of Technology.
68. Feller,W. (1967) In: *An Introduction to Probability Theory and its Applications*. 3d edn. Wiley, NY.
69. Wu,L.-F., Liu,Z., Roberts,S.J., Su,M., Szostak,J.W. and Sutherland,J.D. (2022) Template-Free Assembly of Functional RNAs by Loop-Closing Ligation. *J. Am. Chem. Soc.*, **144**, 13920-13927.
70. Zolaktaf,S., Dannenberg,F., Rudelis,X., Condon,A., Schaeffer,J.M., Schmidt,M., Thachuk,C. and Winfree,E. (2017) In: Brijder, R. and Qian,L. (eds). *DNA Computing and Molecular Programming*. Springer International Publishing, Cham, pp. 172–187.