JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
Cite This: J. Am. Chem. Soc. XXXX, XXX, XXX−XXX
pubs.acs.org/JACS

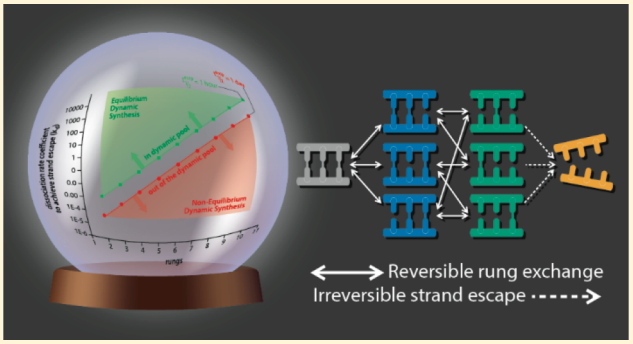
Quantifying Error Correction through a Rule-Based Model of Strand Escape from an [*n*]-Rung Ladder

Morgan M. Cencer,[†] Andrew J. Greenlee,[†] and Jeffrey S. Moore*[,‡,†]

[†]Department of Chemistry, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States
[‡]Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States

Supporting Information

ABSTRACT: The rational design of 3D structures (MOFs, COFs, etc.) is presently limited by our understanding of how the molecular constituents assemble. The common approach of using reversible interactions (covalent or noncovalent) becomes challenging, especially when the target is made from multivalent building blocks and/or under conditions of slow exchange, as kinetic traps and nonequilibrium product distributions are possible. Modeling the time course of the assembly process is difficult because the reaction networks include many possible pathways and intermediates. Here we show that rule-based kinetic simulations efficiently model dynamic reactions involving multivalent building blocks. We studied "strand escape from an [*n*]-rung ladder" as an example of a dynamic process characterized by a complex reaction network. The strand escape problem is important in that it predicts the time a dynamic system needs to backtrack from errors involving [*n*]-misconnections. We quantify the time needed for error correction as a function of the dissociation rate coefficient, strand valency, and seed species. We discuss the simulation results in relation to a simple probabilistic framework that captures the power law dependence on the strand's valency, and the inverse relationship to the rung-opening rate coefficient. The model also tests the synthetic utility of a one-rung (i.e., hairpin) seed species, which, at intermediate times, bifurcates to a long-lived, fully formed [*n*]-rung ladder and a pair of separated strands. Rule-based models thus give guidance to the planning of a dynamic covalent synthesis by predicting time to maximum yield of persistent intermediates for a particular set of rate coefficients and valency.



INTRODUCTION

Dynamic synthesis targets ordered structures made from constituents that are joined through multiple, reversible bonding sites (i.e., the constituents are multivalent with dynamic bonds). In many cases, thermodynamic modeling accurately predicts the major products since equilibrium is readily achieved.[1−3] As increasingly complex structures are targeted,[4] persistent intermediates arise as kinetic traps,[5−7] stymieing the synthesis.[8,9] Kinetic trapping occurs when the rate of constituent exchange is slower than the time scale to perform the overall reaction in the laboratory.[10] Since dynamic chemistries cover a wide kinetic range (Figure 1), the inability to converge to a thermodynamic minimum is more frequently encountered in systems with low dissociation rates.[11] Additionally, kinetic traps are also common with high-valency constituents,[12−15] owing to the low probability of simultaneously dissociating all of the bonding interactions (i.e., high avidity).[15] In extreme cases, escaping kinetic traps becomes a rate limiting step to error correction for a dynamic covalent synthesis.[16]

Error correction is that part of the synthesis pathway that reverses the placement of a constituent whose pattern of bonding connectivity does not adhere to that of the target. Error correction is accomplished either by removing an improperly placed constituent from the assembling structure or by repositioning it into its proper location.[17,18] Error correction is a well-known and generally applicable concept of dynamic, multivalent chemistries.[19−21] Removal of the misplaced constituent is slower the more bonds that must break and the lower the dissociation rate of the individual bonding interactions.[12] Quantifying how strongly valency, affinity, and bond dissociation rates impact error correction is desired for the rational planning of a dynamic synthesis.[22−24] To the best of our knowledge, there is no quantitative description of error correction in the literature. Here we develop a model for the specific case where error correction involves complete removal of an improperly placed constituent.

Rule-based models (RBMs) handle the construction of complex reaction networks algorithmically, making it possible to simulate the time-evolution of dynamic multivalent

Received: August 19, 2019

DOI: 10.1021/jacs.9b08958
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

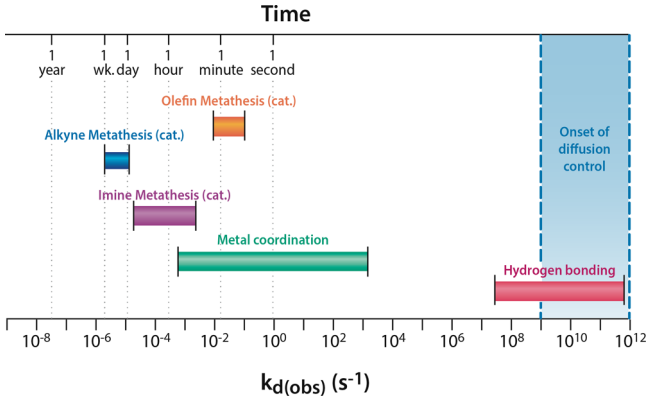


Figure 1. Literature values comparing the dissociation rate coefficients for a variety of reversible reactions (k_d (obs)). Literature references, assumptions, and methods to compare the rate coefficients are provided in the Supporting Information.

reactions.[25] RBMs employ simplifying assumptions that allow a small set of rules to define the reactive chemistry of a system.[26] Rules represent reactions that change the bonds (e.g., bond formation) or state of a reactant (e.g., catalyst deactivation).[27] Once the rules capture the desired attributes of the system, the reaction network is generated and the time evolution is stochastically simulated starting from an initial species (i.e., the seed species). RBMs have been broadly used in biochemistry to handle the complexity of biochemical networks and pathways[28−39] (Table 1). To the best of our knowledge,

Table 1. Applications of Rule-Based Models

topic	system modeled	reference
cellular signaling	oligomerization and phosphorylation of EGFR	Kozer, N., et al.[26]
	protein−protein interactions	Stites, E. C., et al.[27]
	interactions between $\alpha V\beta 3$ integrin and VEGFR2	Montgomery, R., et al.[28]
	FcεRI signaling network	Chylek, L. A., et al.[29]
	cell fate decision-making in the p53 system	Hat, B., et al.[30]
	intracellular competition for calcium	Antunes, G., et al.[31]
DNA repair and tiles	DNA damage repair by nonhomologous end joining	Dolan, D., et al.[32]
	DNA tile self-assembly	Amarioarei, A., et al.[33]
viral capsid assembly	dynamic pathways for T1 capsids	Hagan, M., et al.[34]
	icosahedral viral capsids	Schwartz, R., et al.[35]
other	thermochemical routes to biomass conversion	Rangarajan, S., et al.[36]
	biome mercury exchange	Hartman, J., et al.[37]
dynamic multivalent chemistry	strand escape from an [n]-rung ladder	this work

RBMs have not previously been used to simulate process that are relevant to dynamic chemistry. As shown below, we have found that RBMs are readily adapted to dynamic multivalent chemistry to construct the reaction network and simulate the time course of the system as it progresses toward equilibrium.

Multivalent reaction networks are combinatorially complex.[11] As the valency of the building block increases, the number of reactions and intermediates increases exponentially (see Table 2 for a specific example). Kinetic models of dynamic multivalent reactions must incorporate the many intermediates and pathways relevant to the problem at hand.[40] Time course simulations are then able to generate non-equilibrium product distributions and reveal possible kinetic traps.

Table 2. Characteristics of Reaction Networks for Strand Escape from [n]-Rung Ladders

[n][a]	no. of species[b]	no. of reactions[c]	no. of rules[d]
[2]	4	6	4
[3]	8	21	6
[4]	16	60	8
[5]	32	155	10
[6]	64	378	12
[7]	128	889	14
[8]	256	2040	16
[9]	512	4599	18
[10]	1024	10 230	20

[a][n] is the number of rungs in a double-stranded ladder; more generally, [n] describes the valency of the constituents. [b]Species are defined as ladder (starting reactant), plus all unique intermediates, plus the free strand (product). [c]Reactions are the elementary steps, i.e., a single rung-opening or rung-closing step. Species are connected by these single elementary reactions. Strands are irreversibly separated once the last remaining rung opens. [d]The number of rules required to generate the complete set of reactions listed in the previous column.

In this paper we address the general class of reactions that we refer to as "strand escape from an [n]-rung ladder". The fundamental significance of this reaction class is that it generalizes the dissociation of a pair of multivalent constituents, valency being equated to the number of rungs, [n]. The specific chemistry does not matter here, as this model is applicable to any type of dynamic interaction. The kinetics of strand escape predict the time required for error correction (i.e., constituent removal) in which backtracking requires the simultaneous disconnection of [n] interactions. Table 2 shows that the number of species and number of elementary reactions associated with the reaction network increase exponentially with the number of rungs ([n]) (see SI). By employing a simplifying set of assumptions (vide infra), the number of rules required to construct the network only depends linearly on [n] (Table 2). The RBM approach to model construction enables the simulation of complex reaction networks making it possible to systematically probe factors such as valency, affinity, and bond dissociation rates of dynamic multivalent processes.

MODEL CONSTRUCTION

We developed strand escape from an [n]-rung ladder as a model to quantify the time needed for error correction. While the model refers to [n]-rung ladders, the ladders are equivalent to a set of misconnections between a pair of constituents. For example, escape from a 5-rung ladder is equivalent to the largest possible number of misconnections between two 6-rung strands. It also corresponds to an error in a periodic lattice that requires 5 simultaneous disconnections in order to remove the misplaced constituent (Figure 2, [5]-rung error). We have focused on the disassembly process (i.e., strand escape), rather than the assembly process (i.e., ladder formation), to specify a precise mode by which error correction is achieved (i.e., the

DOI: 10.1021/jacs.9b08958
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

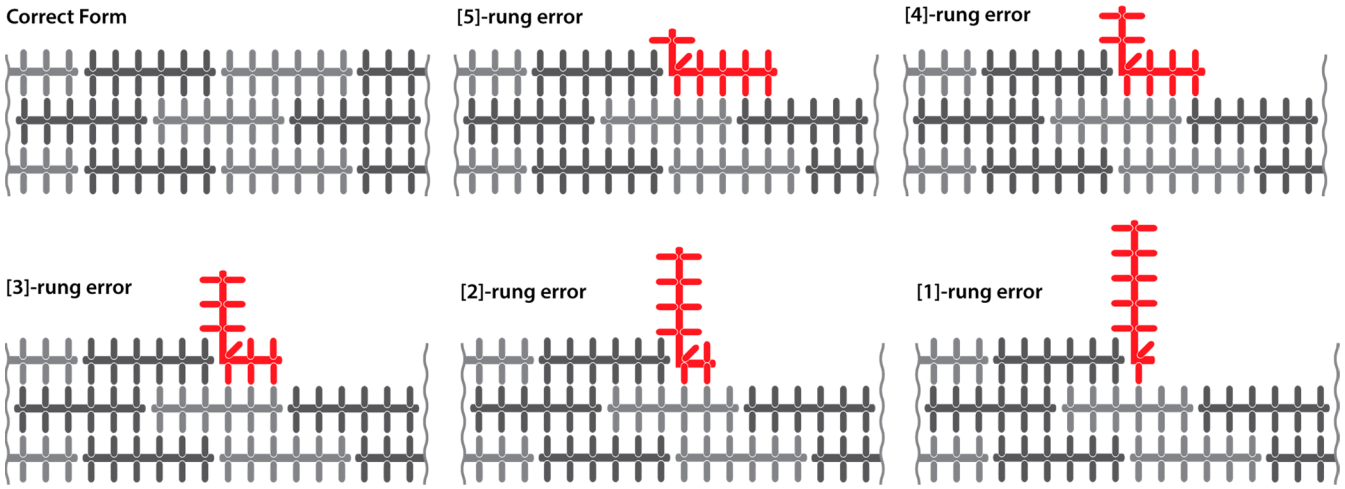


Figure 2. Example of errors in COF formation with [5] through [1] misconnections demonstrating that this work can be expanded to 2D and 3D structures. The misconnected strand in highlighted in red. The light and dark gray shading of the rest of the COF is only provided to guide the eye.

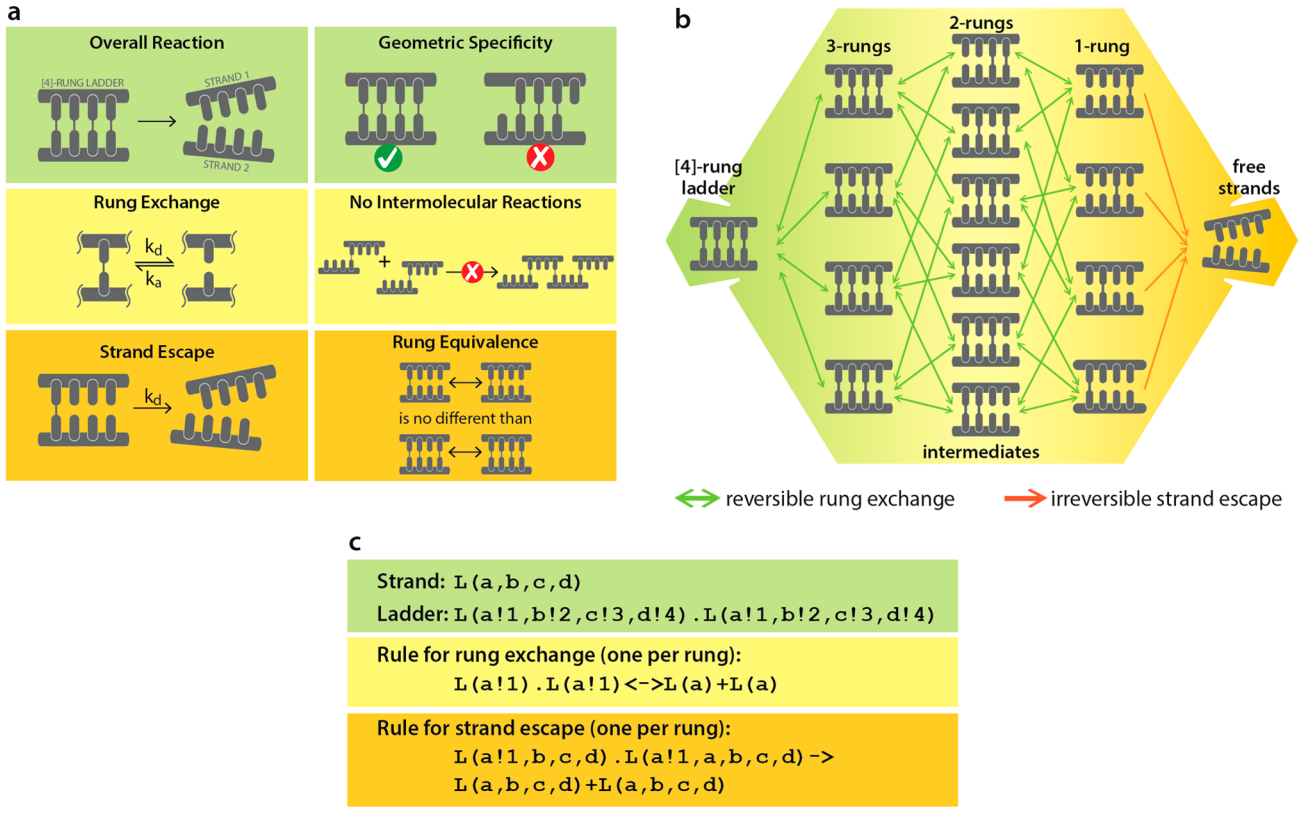


Figure 3. (a) Mechanistic assumptions governing model construction include dynamic rung exchange, strand escape and absence of intermolecular reactions (i.e., infinite dilution case), geometric specificity, and rung equivalence. (b) The network of reactions shows the reversible and irreversible steps occurring in strand escape from the [4]-rung ladder. The network converges on two nodes—the fully bonded ladder and the free strands. The network also shows the number of species (16) and reactions (60). As shown in Table 2, the number of intermediates and reactions grows exponentially with increasing [*n*]. (c) The in-model definitions of free strand and ladder, and the rules used in the model for rung exchange and strand escape. Full text of the model can be found in the SI.

constituent removal mode). Our model is useful for (1) testing if this mode of error correction is the rate limiting step in a dynamic synthesis, and (2) predicting the occurrence of kinetically trapped structures.

Our RBM for [*n*]-rung ladder disassembly is based on the assumptions described in Figure 3a. The Supporting Information provides the details of how these assumptions were implemented into BioNetGen,[41] an RBM simulator. Using only eight rules (Figure 3c), the complete reaction network for strand escape from the [4]-rung ladder was algorithmically generated (Figure 3b). Reaction arrows are drawn between any two species related by the execution of a single elementary step (i.e., rung association (right to left) or rung dissociation (left to right)). Through a sequence of steps, the [4]-rung ladder is transformed to a pair of free strands, the shortest pathway involving four steps and three intermediates

DOI: 10.1021/jacs.9b08958
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

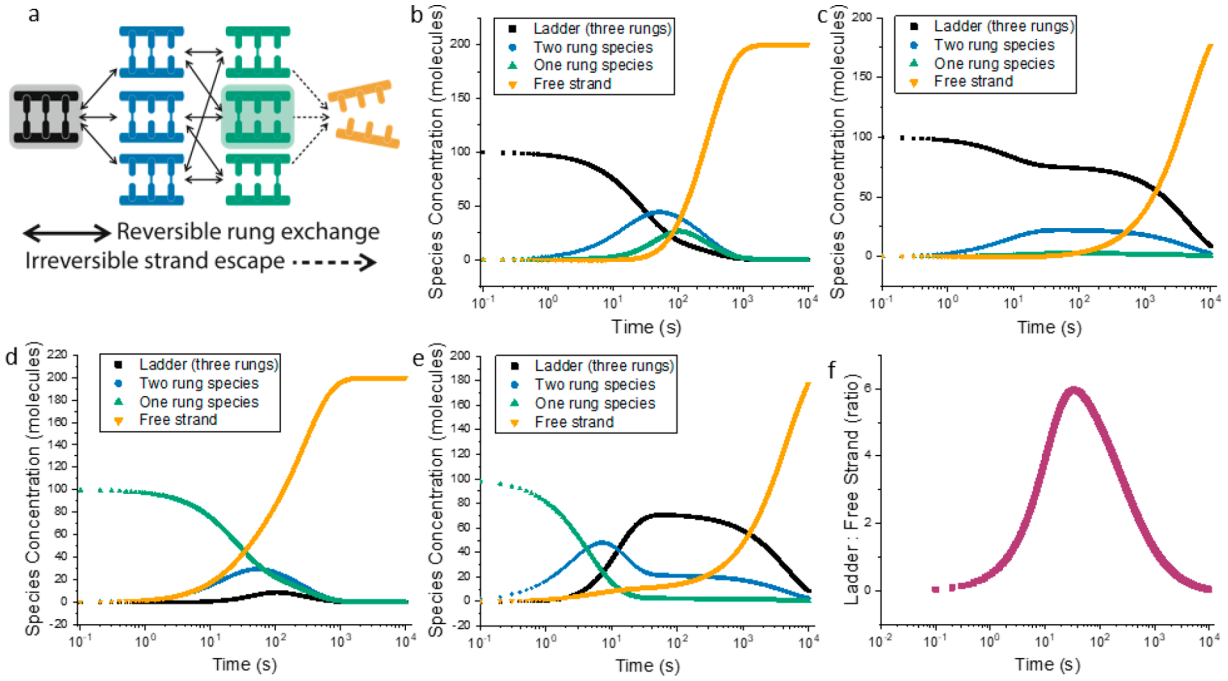


Figure 4. (a) The reaction network for escape from a [3]-rung ladder. (b) The distribution of products for a reaction where the seed species is the three rung ladder (shaded black in (a)), which leads to a rise in two and one rung species as the rungs dynamically open and close. Eventually the free strands separate. In this reaction $k_d = k_a = 0.01\ s^{-1}$, (c) shows the same reaction as (b) except that $k_d = 0.01\ s^{-1}$, while $k_a = 0.1\ s^{-1}$, leading to kinetic trapping of both the ladder species and the two-rung species. The difference between a dynamic intermediate and a kinetically trapped species is the lifetime of the species as compared to the reaction time. (d) The distribution of products over a reaction where $k_d = k_a = 0.01\ s^{-1}$ and the seed species is a hairpin type structure (shaded green in (a)). (e) The same reaction as (d) except that $k_d = 0.01\ s^{-1}$, while $k_a = 0.1\ s^{-1}$, leading to a steady state concentration of ladder, and 2-rung species. (f) The ratio between ladder and free strand from part (e). It peaks at a ratio of 6, at 38 s.

(Figure 3b). All the other reaction networks for ladders with [2]- to [10]-rungs were similarly generated, as the assumptions are general for any length ladder. Each rung is assumed to independently undergo reversible opening (rung dissociation) and closing (rung association). The rates of these intramolecular reactions are governed by first order rate coefficients k_d and k_a, respectively. Cooperativity in rung binding was not considered; in this work the rate coefficients for opening and closing bonds are assumed to be independent of local environment. Cooperativity (either positive or negative) is possible by adding rules that define state dependent rates (e.g., the rate depends on the bonding state of adjacent rungs). Strand escape occurs irreversibly when the final rung dissociates. The infinite dilution case allows the model to focus solely and specifically on the escape step of error correction. The rate coefficient for this step is also assumed to be k_d. Since the model assumes that strand escape takes place under conditions of infinite dilution, all intermolecular interactions are prohibited. For simplicity, we only consider the case where the set of rungs on one strand bond to a contiguous set of rungs on the opposite strand without rung crossover. The resulting reaction network (Figure 3b) converges at two places corresponding to (1) the fully connected [4]-rung ladder and (2) the pair of separated strands. As will become apparent, these points of convergence are the major species that accumulate transiently and as equilibrium is approached. We surmise that convergent points on reaction networks are generally either the final equilibrium product(s) or they are intermediates that become kinetically trapped (i.e., long-lived, persistent intermediates). Beyond [4]-rungs, these multivalent reaction networks are difficult to visualize, which in general, makes predicting reaction outcomes and identifying the location of traps more challenging. Using these models, we demonstrate the utility of RBMs for validating synthetic design starting from an advanced intermediate and quantify the time required for error correction (i.e., constituent removal).

MODELING THE TIME COURSE OF STRAND ESCAPE

Once the networks are generated, the time course of the overall reaction progress is repeatedly simulated using sets of values for k_a and k_d. Figure 4 shows simulation results for two different kinetic values starting from the fully formed [3]-rung ladder. When the rate coefficient for rung breaking is equal to the rate coefficient for rung making, the [3]-rung ladder (shaded black in Figure 4a) gradually disappears, passing through the expected [2]-rung and, [1]-rung intermediates on its way to free the strands (Figure 4b, $k_d = k_a$). When the kinetic coefficients favor rung formation (i.e., $k_d < k_a$) there are distinctive changes in the time dependent concentration profiles (Figure 4c). Noteworthy are the occurrence of broad plateaus that persist for a significant period of time but eventually fade as the free strands are generated. These plateaus are signatures of kinetic traps. Kinetic traps are persistent, long-lived intermediates, where the "long" is defined by the time scale of the experiment.

The formation of a kinetic trap is more noticeable when starting from an advanced intermediate. We simulated escape from a [3]-rung ladder where the starting (i.e., seed) species was a "hairpin" (shaded green in Figure 4a). When the rate coefficient for rung breaking is equal to the rate coefficient for

DOI: 10.1021/jacs.9b08958
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

rung making ($k_a = k_d$), all the intermediates on the reaction network are still reached, including a transient appearance of the fully bonded [3]-rung ladder (Figure 4d). When the kinetic coefficients favor rung formation (i.e., $k_d < k_a$) the ladder species becomes kinetically trapped for a significant period of time (Figure 4e). These trapped species and their time dependence are of particular interest to the experimental chemist. For instance, the [3]-rung ladder in Figure 4e might represent a defect in an assembly process whose "error-correction mechanism" requires backtracking to the free strands. To gauge the likelihood of a successful dynamic multivalent synthesis, the chemist would need to know the time to undo the defect (i.e., error correction via constituent removal). In another instance, the [3]-rung ladder in Figure 4e might represent the desired kinetic product for an assembly process that begins from the hairpin. Synthetic planning will benefit from knowing the time point that maximizes the yield of desired product (i.e., the [3]-rung ladder) in addition to the time point that maximizes the ratio (Figure 4f) of the desired product relative to irreversible losses (i.e., free strands). In the reaction shown in Figure 4e, the largest ratio for fully bonded ladder to free strand is found at 38 s (Figure 4f), but the maximum yield of ladder is found at 61 s. Which of these maxima is the best time to stop a reaction depends on the ease of separation of the mixture and the potential to isolate and subsequently resubject intermediates to the dynamic conditions in order to recover a second crop of desired product.

To expand on the synthetic example, we systematically varied the rate coefficient ratio for rung dissociation (k_d) and rung formation (k_a) starting from the hairpin species leading to ladders having [2]- to [5]-rungs. By varying the rate coefficient ratio systematically for various length ladders, we are able to analyze the effect of kinetic coefficients and the effect of valency on reaction progression and kinetic trapping. The ladder lifetime (defined in Figure 5a) decreases as K_D increases ($K_D = k_d/k_a$). For $K_D < 1$, the ladder lifetime depends strongly on valency (Figure 5b) while the maximum yield exhibits only a slight dependence on valency (Figure 5c). Interestingly, for a given value of K_D, the time to reach maximum ladder yield is markedly dependent on valency (Figure 5d). These results show that a thorough understanding of the time evolution of a system is necessary to capture the desired nonequilibrium product distribution.

QUANTIFYING THE TIME REQUIRED FOR ERROR CORRECTION

To expand on the error-correction model for a dynamic multivalent synthesis that targets an equilibrium product, we systematically examined how strand escape times of 1 h and 1 day depend on rung dissociation rate coefficient and valency (Figure 6). The plot in Figure 6 is for a case that is likely to result in kinetic traps (i.e., where the rung exchange equilibrium favors rung formation ($K_D = 0.1$)) (results for K_D equal to 1 and 10 are shown in SI Figure S4). Times of 1 h and 1 day were selected to represent error-correction times that are of significance to a dynamic multivalent synthesis. Times that are shorter than 1 h are characteristic of [n]-rung species that remain in the "dynamic pool" on a reasonable laboratory time frame. In contrast, times longer than 1 day are characteristic of [n]-rung species that are in the "static pool" on a reasonable laboratory time frame. Error correction is seen to become unlikely for high-valent systems or systems with small dissociation rate coefficients. To achieve an equilibrium target

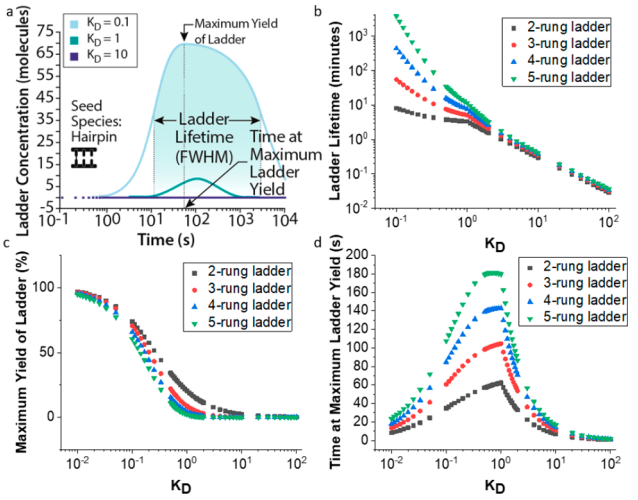


Figure 5. Different behavior is seen with different rate coefficients. In all cases shown here the seed species was a hairpin (i.e., one rung is closed on one end of the strands while the remaining rungs are open, shown in (a) in schematic form). Note that $K_D = k_d/k_a$. For $K_D < 1$, k_a was held constant at 0.01 s^{-1} while k_d was varied. For $K_D > 1$, k_d was held constant at 0.01 s^{-1} while k_a was varied. (a) The concentration of [3]-ladder for three different K_D when the seed species is a hairpin. When k_a is greater than or equal to k_d there is a distinct time domain during which the [n]-rung ladder formation exhibits steady state behavior, known as the ladder lifetime. Ladder lifetime is defined as the full width at half-maximum height (fwhm) for the ladder concentration peak. (b) The effects of K_D and rung number on the ladder lifetime when the starting species is a hairpin. When k_a is greater than k_d, ladder lifetimes increase, with the number of rungs having a strong effect on the amount of increase. When k_a is less than k_d, all ladder lifetimes converge toward zero with increasing k_d. (c) The effects of K_D and rung number on the maximum percent yield of ladder when the starting species is a hairpin. Increasing the number of rungs decreases overall yield, although the change between [n] rungs and [n + 1] rungs decreases with increasing [n]. (d) The effects of rates and rung number on the time at which maximum ladder yield is reached when the starting species is a hairpin. Time at maximum ladder concentration decreases with large k_d due to limited formation of ladder. The same trend is seen with large k_a due to rapid formation of ladder and minimal separation into free strand.

using a dynamic multivalent synthesis involving a building block having a particular valency, the kinetic coefficient needed for all species to remain in the dynamic pool may limit the choice of viable dynamic chemistries according to the ranges specified in Figure 1.

The relationship between number of rungs, ladder escape half-life, and the rung exchange rate coefficients are reasoned with a simple probabilistic framework (eq 1, full derivation in the SI) that predicts useful relationships between the valency of the species ([n]), the kinetic parameters (k_a and k_d), and the strand escape half-life ($t_{1/2}^{escp}$, i.e., the time at which half of the strands have escaped the ladder architecture).

$$t_{1/2}^{escp} = \left(\frac{1 + K_D}{K_D}\right)^{[n]} \times \frac{\ln(2)}{k_d} \quad (1)$$

This equation predicts a power law relationship between $t_{1/2}^{escp}$ and the valency of the species, consistent with the results plotted in Figure 6. The equation also predicts an inverse relationship between ladder escape half-life and k_d, which we have confirmed in simulation (SI Figures S2 and S3). The dependency of $t_{1/2}^{escp}$ on [n] and k_d is evident in Figure 6. Half-

DOI: 10.1021/jacs.9b08958
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

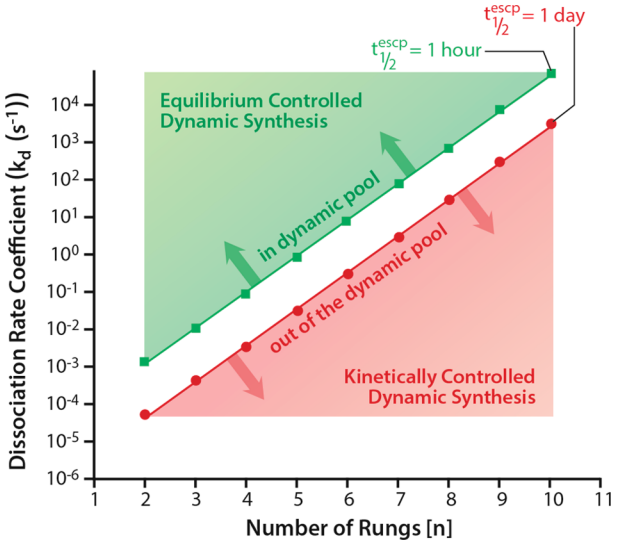


Figure 6. Error correction profile for K_D = 0.1 (an example that generates kinetic traps). The model predicts the dissociation rate coefficient (k_d) needed to achieve specific ladder escape half-lives $t_{1/2}^{escp}$ based on the number of rungs in the ladder. The dissociation rate coefficient (k_d) needed to achieve a $t_{1/2}^{escp}$ of 1 h or less is shaded green while the dissociation rate coefficient (k_d) needed to achieve a $t_{1/2}^{escp}$ of 1 day or more is shaded red. Additional information is in SI Figures S3 and S4. Any species with a $t_{1/2}^{escp}$ of an hour is considered within the dynamic pool and will lead to equilibrium DC synthesis. Any species with a $t_{1/2}^{escp}$ of a day (or more) is outside of the dynamic pool and will lead to nonequilibrium DC synthesis and kinetic trapping. Interesting behavior and trapping are also found in the region between those two times.

lives are highly dependent on both the valency and the rate coefficients. Accounting for the impact of both valency and kinetics is critical for rationally designing a dynamic multivalent synthesis.

Imine ladder formation provides a qualitative example of how these findings are applied to an experimental system. In Figure 1, we see that the fastest imine metathesis has a k_d of about 10^{-3} s^{-1}. With this k_d we predict that three rung ladders are likely to experience some trapping as their escape half-life is longer than 1 h, and that four or more rungs are likely to see significant kinetic trapping as escape half-lives for these species are 1 day or longer (Figure 6). Experimentally, it has been found that imine ladders of four or more rungs are indeed kinetically trapped.[42] Raising the reaction temperature notably increased the amount of strand scrambling. By raising the temperature, the value of k_d shifts to a larger value, returning misconnected strands back into the dynamic pool. Experimental determination of bond dissociation rates can thus establish the likelihood that a system will form kinetic traps. Using our model along with experimentally obtained kinetic details will help guide the researcher about the conditions for which error correction is rapid, thus avoiding lengthy optimization studies.

CONCLUSIONS

The ladder disassembly model demonstrates the utility of RBMs for dynamic multivalent chemistry while quantitatively simulating the dynamics of error correction via constituent removal for a reversible, multivalent system. The valency of the constituent, seed species, and kinetic parameters of the system are all critical to whether the rate of constituent removal is fast enough to allow backtracking as needed to correct errors. By combining experimentally determined kinetic parameters and RBMs, we believe that chemists will be able to rationally design increasingly complicated synthetic targets. While this publication has focused on 1D [n]-rung ladder misconnection types, it is possible to extend this approach to more complex structures such as 2D materials and 3D MOFs and COFs. We are currently working to combine experimentally derived kinetic coefficients and RBMs to rationally design advanced constructs. Future models will include intermolecular reactions to describe the complete synthetic pathway, including the off-target pathways (i.e., error formation and reversal).

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.9b08958.

Simulation files and information, additional graphs, full derivation of the probabilistic framework presented, and explanation of the data in Figure 1 (PDF)

AUTHOR INFORMATION

Corresponding Author

*jsmoore@illinois.edu

ORCID

Morgan M. Cencer: 0000-0003-2806-8317
Andrew J. Greenlee: 0000-0002-9085-6309
Jeffrey S. Moore: 0000-0001-5841-6269

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. CHE 16-10328 and Grant No. CHE 19-04180. We are deeply grateful to Dr. Chris Pattillo, Summer Laffoon, and Professor Corrin Clarkson for valuable discussions.

REFERENCES

(1) Belowich, M. E.; Stoddart, J. F. Dynamic Imine Chemistry. *Chem. Soc. Rev.* **2012**, *41* (6), 2003−2024.

(2) Cao, N.; Wang, Y.; Zheng, X.; Jiao, T.; Li, H. Controllable Self-Assembly of Pills and Cages via Imine Condensation for Silver Cation Detection. *Org. Lett.* **2018**, *20*, 7447−7450.

(3) Giuseppone, N.; Schmitt, J. L.; Schwartz, E.; Lehn, J. M. Scandium(III) Catalysis of Transimination Reactions. Independent and Constitutionally Coupled Reversible Processes. *J. Am. Chem. Soc.* **2005**, *127* (15), 5528−5539.

(4) Beuerle, F.; Gole, B. Covalent Organic Frameworks and Cage Compounds: Design and Applications of Polymeric and Discrete Organic Scaffolds. *Angew. Chem., Int. Ed.* **2018**, *57* (18), 4850−4878.

(5) Baram, J.; Weissman, H.; Rybtchinski, B. Supramolecular Polymer Transformation: A Kinetic Study. *J. Phys. Chem. B* **2014**, *118* (41), 12068−12073.

(6) Grötsch, R. K.; Wanzke, C.; Speckbacher, M.; Angı, A.; Rieger, B.; Boekhoven, J. Pathway Dependence in the Fuel-Driven Dissipative Self-Assembly of Nanoparticles. *J. Am. Chem. Soc.* **2019**, *141* (25), 9872−9878.

(7) Zhu, G.; Liu, Y.; Flores, L.; Lee, Z. R.; Jones, C. W.; Dixon, D. A.; Sholl, D. S.; Lively, R. P. Formation Mechanisms and Defect Engineering of Imine-Based Porous Organic Cages. *Chem. Mater.* **2018**, *30* (1), 262−272.

DOI: 10.1021/jacs.9b08958
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

(8) Wang, Q.; Yu, C.; Zhang, C.; Long, H.; Azarnoush, S.; Jin, Y.; Zhang, W. Dynamic Covalent Synthesis of Aryleneethynylene Cages through Alkyne Metathesis: Dimer, Tetramer, or Interlocked Complex? *Chem. Sci.* **2016**, *7* (5), 3370−3376.
(9) Mattia, E.; Otto, S. Supramolecular Systems Chemistry. *Nat. Nanotechnol.* **2015**, *10* (2), 111−119.
(10) Ding, H.; Chen, R.; Wang, C. Organic Cages through Dynamic Covalent Reactions. *Dyn. Covalent Chem. React. Princ. Appl.* **2017**, 165−205.
(11) Jin, Y.; Yu, C.; Denman, R. J.; Zhang, W. Recent Advances in Dynamic Covalent Chemistry. *Chem. Soc. Rev.* **2013**, *42* (16), 6634−6654.
(12) Hartley, C. S.; Elliott, E. L.; Moore, J. S. Covalent Assembly of Molecular Ladders. *J. Am. Chem. Soc.* **2007**, *129* (15), 4512−4513.
(13) Wei, T.; Furgal, J. C.; Jung, J. H.; Scott, T. F. Long, Self-Assembled Molecular Ladders by Cooperative Dynamic Covalent Reactions. *Polym. Chem.* **2017**, *8* (3), 520−527.
(14) Fasting, C.; Schalley, C. A.; Weber, M.; Seitz, O.; Hecht, S.; Koksch, B.; Dernedde, J.; Graf, C.; Knapp, E. W.; Haag, R. Multivalency as a Chemical Organization and Action Principle. *Angew. Chem., Int. Ed.* **2012**, *51* (42), 10472−10498.
(15) Tjandra, K. C.; Thordarson, P. Multivalency in Drug Delivery − When Is It Too Much of a Good Thing? *Bioconjugate Chem.* **2019**, *30* (3), 503−514.
(16) Jin, Y.; Wang, Q.; Taynton, P.; Zhang, W. Dynamic Covalent Chemistry Approaches toward Macrocycles, Molecular Cages, and Polymers. *Acc. Chem. Res.* **2014**, *47* (5), 1575−1586.
(17) Whitesides, G. M.; Mathias, J. P.; Seto, C. T. Molecular Self-Assembly and Nanochemistry: A Chemical Strategy for the Synthesis of Nanostructures. *Science (Washington, DC, U. S.)* **1991**, *254*, 1312−1319.
(18) Lehn, J. M. Toward Self-Organization and Complex Matter. *Science (Washington, DC, U. S.)* **2002**, *295* (5564), 2400−2403.
(19) Hasell, T.; Cooper, A. I. Porous Organic Cages: Soluble, Modular and Molecular Pores. *Nat. Rev. Mater.* **2016**, DOI: 10.1038/natrevmats.2016.53.
(20) Acharyya, K.; Mukherjee, P. S. Organic Imine Cages: Molecular Marriage and Applications. *Angew. Chem., Int. Ed.* **2019**, *58*, 8640.
(21) Rowan, S. J.; Cantrill, S. J.; Cousins, G. R. L.; Sanders, J. K. M.; Stoddart, J. F. Dynamic Covalent Chemistry. *Angew. Chem., Int. Ed.* **2002**, *41*, 1460.
(22) Komáromy, D.; Stuart, M. C. A.; Monreal Santiago, G.; Tezcan, M.; Krasnikov, V. V.; Otto, S. Self-Assembly Can Direct Dynamic Covalent Bond Formation toward Diversity or Specificity. *J. Am. Chem. Soc.* **2017**, *139* (17), 6234−6241.
(23) Mahon, C. S.; Fulton, D. A. Mimicking Nature with Synthetic Macromolecules Capable of Recognition. *Nat. Chem.* **2014**, *6* (8), 665−672.
(24) Panettieri, S.; Ulijn, R. V. Energy Landscaping in Supramolecular Materials. *Curr. Opin. Struct. Biol.* **2018**, *51*, 9−18.
(25) Chylek, L. A.; Harris, L. A.; Tung, C.-S.; Faeder, J. R.; Lopez, C. F.; Hlavacek, W. S. Rule-Based Modeling: A Computational Approach for Studying Biomolecular Site Dynamics in Cell Signaling Systems. *WIREs Syst. Biol. Med.* **2014**, *6* (1), 50.
(26) Chylek, L. A.; Harris, L. A.; Faeder, J. R.; Hlavacek, W. S. Modeling for (Physical) Biologists: An Introduction to the Rule-Based Approach. *Phys. Biol.* **2015**, *12* (4), 045007.
(27) Hlavacek, W. S.; Faeder, J. R.; Blinov, M. L.; Posner, R. G.; Hucka, M.; Fontana, W. Rules for Modeling Signal-Transduction Systems. *Sci. Signaling* **2006**, *2006* (34), 1−18.
(28) Kozer, N.; Barua, D.; Orchard, S.; Nice, E. C.; Burgess, A. W.; Hlavacek, W. S.; Clayton, A. H. A. Exploring Higher-Order EGFR Oligomerisation and Phosphorylation - A Combined Experimental and Theoretical Approach. *Mol. BioSyst.* **2013**, *9* (7), 1849−1863.
(29) Stites, E. C.; Aziz, M.; Creamer, M. S.; Von Hoff, D. D.; Posner, R. G.; Hlavacek, W. S. Use of Mechanistic Models to Integrate and Analyze Multiple Proteomic Datasets. *Biophys. J.* **2015**, *108* (7), 1819−1829.
(30) Rangarajan, S.; Bhan, A.; Daoutidis, P. Rule-Based Generation of Thermochemical Routes to Biomass Conversion. *Ind. Eng. Chem. Res.* **2010**, *49* (21), 10459−10470.
(31) Hartman, J. S.; Weisberg, P. J.; Pillai, R.; Ericksen, J. A.; Kuiken, T.; Lindberg, S. E.; Zhang, H.; Rytuba, J. J.; Gustin, M. S. Application of a Rule-Based Model to Estimate Mercury Exchange for Three Background Biomes in the Continental United States. *Environ. Sci. Technol.* **2009**, *43* (13), 4989−4994.
(32) Bazzazi, H.; Zhang, Y.; Jafarnejad, M.; Popel, A. S. Computational Modeling of synergistic interactions between aVb3 integrin and VEGFR2 in endothelial cells: Implications for the mechanism of action of angiogenesis-modulating integrin-binding peptides. *J. Theor. Biol.* **2018**, *455* (5), 212−221.
(33) Chylek, L. A.; Holowka, D. A.; Baird, B. A.; Hlavacek, W. S. An Interaction Library for the FcεRI Signaling Network. *Front. Immunol.* **2014**, *5*, 1−16.
(34) Hat, B.; Kochańczyk, M.; Bogdał, M. N.; Lipniacki, T. Feedbacks, Bifurcations, and Cell Fate Decision-Making in the P53 System. *PLoS Comput. Biol.* **2016**, *12* (2), 1−28.
(35) Antunes, G.; Roque, A. C.; Simoes De Souza, F. M. Modelling Intracellular Competition for Calcium: Kinetic and Thermodynamic Control of Different Molecular Modes of Signal Decoding. *Sci. Rep.* **2016**, *6*, 1−12.
(36) Dolan, D. W. P.; Zupanic, A.; Nelson, G.; Hall, P.; Miwa, S.; Kirkwood, T. B. L.; Shanley, D. P. Integrated Stochastic Model of DNA Damage Repair by Non-Homologous End Joining and P53/P21- Mediated Early Senescence Signalling. *PLoS Comput. Biol.* **2015**, *11* (5), 1−19.
(37) Amarioarei, A.; Spencer, F.; Itscus, C.; Tusa, I.; Dobre, A.; Barad, G.; Trandafir, R.; Paun, M.; Czeizler, E. Computational Approaches for the Programmed Assembly of Nanocellulose Meshes. *Fed. Log. Conf.* **2018**, 1−5.
(38) Hagan, M. F.; Chandler, D. Dynamic Pathways for Viral Capsid Assembly. *Biophys. J.* **2006**, *91* (1), 42−54.
(39) Schwartz, R.; Shor, P. W.; Prevelige, P. E.; Berger, B. Local Rules Simulation of the Kinetics of Virus Capsid Self-Assembly. *Biophys. J.* **1998**, *75* (6), 2626−2636.
(40) Markvoort, A. J.; Ten Eikelder, H. M. M.; Hilbers, P. A. J.; De Greef, T. F. A. Fragmentation and Coagulation in Supramolecular (Co)Polymerization Kinetics. *ACS Cent. Sci.* **2016**, *2* (4), 232−241.
(41) Arora, A.; Sheehan, R. P.; Hogg, J. S.; Faeder, J. R.; Korsunsky, I.; Sekar, J. A. P.; Gupta, S.; Harris, L. A.; Tapia, J.-J.; Barua, D. BioNetGen 2.2: Advances in Rule-Based Modeling. *Bioinformatics* **2016**, *32* (21), 3366−3368.
(42) Elliott, E. L.; Hartley, C. S.; Moore, J. S. Covalent Ladder Formation Becomes Kinetically Trapped beyond Four Rungs. *Chem. Commun.* **2011**, *47* (17), 5028−5030.

DOI: 10.1021/jacs.9b08958
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX