

# Kinetic and Sequence Regulation in Copolymerization Through Emergent Reactant Ordering

Ryan Lyle Hamblin  
Albuquerque, New Mexico, USA

B.S. Physics, University of New Mexico, 2019

A Dissertation presented to the Graduate Faculty of the University of Virginia in  
Candidacy for the Degree of Doctor of Philosophy

Department of Chemistry  
University of Virginia  
May 2024

# Kinetic and Sequence Regulation in Copolymerization Through Emergent Reactant Ordering

Ryan Lyle Hamblin

Advisor: Professor Kateri H. DuBay

Committee Members:

Professor Sergei A. Egorov, Professor Andreas Gahlman, and Professor Leonid V. Zhigilei

Sequence control in synthetic copolymers remains a tantalizing objective in polymer science, as the phase behavior, self-organization, and bulk material properties of copolymers are intrinsically dependent on their primary comonomer sequences. Achieving precise control over monomer sequence in synthetic copolymerizations is challenging, as sequence determination is influenced not only by the reaction conditions and the properties of the reactants, but also by the statistical nature of the copolymerization process itself. Further, characterizing the primary sequence of a synthetic copolymer is a significant challenge, making the experimental study of sequence development intractable with current methods. Despite these difficulties, greater understanding of sequence development throughout the polymerization process will aid the design of simple, generalizable methods to control sequence and tune supramolecular assembly. To this end, this dissertation utilizes a reactive, Langevin dynamics model of copolymerization to directly observe sequence development *in silico* throughout the reaction. This allows for direct comparison to standard statistical theories of copolymerization processes, as well as direct control over system parameters which may influence the reaction. We particularly target conditions in which reaction driven phase change behaviors occur, which lead to reactant organization and emergent heterogeneity not accounted for in traditional theories. We find that differences in non-bonded attraction strengths between comonomers on the order of thermal fluctuations drive a reactant assembly process, leading to a shift in reaction kinetics, molecular weight distribution, and primary sequence that is not captured by Mayo-Lewis and Flory-Schulz theories. We further explore how differences in solvent selectivity may give rise to such self-organization, leading to sequence biasing and the formation of polymer structures with a wide range of morphologies and composition distributions. Additionally, we examine the influence of chain stiffness in concert with these self-assembly behaviors, exploring a transition to nematic ordering that occurs for oligomers of sufficient persistence length. Such liquid crystalline ordering introduces a characteristic length scale into the system, which

both significantly enriches the formation of specific chain and sequence repeat lengths and shifts in response to the relative diffusive and reactive timescales within the system. This work provides new fundamental insights into the impacts of collective and emergent reactant behaviors on the kinetics and sequence development of copolymers which are not captured by current theories. It develops provides greater understanding of reaction conditions which produce particular sequence behaviors, allowing for the informed design of a sequence-controlled copolymerization in a simple one-pot synthetic method.

## Acknowledgements

This endeavor would not have been possible without my advisor, Kateri DuBay, for her creative insight, careful problem solving, dedication, and boundless enthusiasm and patience, from which I have learned so much. Thank you for continuing to support me in my career and in life, I owe you a great debt of gratitude. I am also grateful for the efforts of my committee and the faculty at UVA, who helped to guide me along this path; to Andreas Gahlmann, Sergei Egorov, Leonid Zhigilei, and Eric Herbst, you have my thanks. There are countless others whose guidance and mentorship I have relied on in my academic journey, but particular thanks are owed to Stacy Copp for her encouragement and insight while overseeing my first forays in the laboratory, and for her continued support of my career. I am a better scientist today for your efforts. I also wish to thank Gabriel Montaña and Sergei Ivanov for providing their expertise and support to my early research efforts, James Thomas and Ivan Deutsch for their engaging and rigorous introduction to foundational theories of statistical and quantum mechanics, and Kevin Gant for cultivating a life-long interest in the physical sciences. I also owe thanks to those who worked beside me in my research group: Zhongmin Zhang, Wenxin Xu, and Nhu Nguyen for the engaging discussions on polymer behavior, Steven Merz and Jacob Swartley for sharpening my coding skills, and Jess Niblo for your scientific insight and valuable perspective, and for your friendship, encouragement, and advice. To all of you, I hope to be able to pay forward the mentorship and guidance, and to share with others the curiosity and wonder of the natural world, which you have shown me.

I also owe thanks to friends and family, too numerous to name, for their love, support, and patience while I have undertaken this effort. A particular thanks to my parents, Randolph and Pamela, and to each of my siblings Amanda, Russell, Jennie, and Jessica, for their unshakable belief in me regardless of my belief in myself. There are too many ways to name that you all have shaped my life, and this never would have happened without you. Thank you.

Finally, thank you, dear reader, for your interest in this work. I hope you find some piece of knowledge within that serves you in some way.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Synthetic copolymers and sequence control . . . . .	1
1.2	Copolymer sequence influences material properties and behaviors . . . . .	2
1.3	Sequence influencing effects in copolymerization . . . . .	3
1.4	Dissertation overview . . . . .	4
<b>2</b>	<b>Methods</b>	<b>7</b>
2.1	Model details . . . . .	7
2.1.1	Monomer structure and interactions . . . . .	7
2.1.2	Symmetric and asymmetric effective inter-monomer interactions . . . . .	10
2.1.3	Radial and angular harmonic potentials describe bonded interactions . . . . .	12
2.2	Simulation progression and bond formation . . . . .	13
2.2.1	Sampling via Langevin dynamics . . . . .	13
2.2.2	Modeling a step-growth reaction mechanism . . . . .	13
2.2.3	Reaction activation energies . . . . .	14
2.3	System descriptors and simulation analysis metrics . . . . .	15
2.3.1	Chain length and block length statistics . . . . .	15
2.3.2	Quantifying deviations from Flory-Schulz and Markovian predictions . . . . .	18
2.3.3	Chain length dependence of bonding statistics . . . . .	18
2.3.4	Quantifying chain stiffness and nematic alignment . . . . .	19
<b>3</b>	<b>Collective Reactant Behaviors and Kinetic Impacts</b>	<b>22</b>
3.1	Introduction . . . . .	22
3.2	Ideal second-order kinetics of step-growth copolymerization . . . . .	24
3.3	Local sequence statistics depend on non-bonded attractions, activation energies, and emergent assemblies . . . . .	34

3.4	Differing attraction strengths and activation energies can drive long-range sequence biasing . . . . .	39
3.5	Activation energies and non-bonded attractions influence block and chain length distributions . . . . .	42
3.6	Summary and conclusions . . . . .	49
<b>4</b>	<b>Reactant Assembly In Selective Solvent Conditions</b>	<b>51</b>
4.1	Introduction . . . . .	51
4.2	Solvent mediated attractions can drive an emergent assembly process, altering reaction kinetics . . . . .	52
4.3	Oligomer sequences depend on solvent selectivity . . . . .	56
4.4	Oligomer persistence length impacts both sequence and aggregate structure	60
4.5	Nascent sequence and selective solvation influence the composition of different regions within the aggregates . . . . .	63
4.6	Summary and conclusions . . . . .	66
<b>5</b>	<b>Nematic Ordering and Emergent Characteristic Block Length</b>	<b>68</b>
5.1	Introduction . . . . .	68
5.2	Reaction conditions shift characteristic block length . . . . .	71
5.3	Block length distribution shifts correspond to changes in relative kinetic timescales . . . . .	73
5.4	Characteristic block length arises from emergent changes in bonding behaviors . . . . .	75
5.5	Oligomer aggregate structures and bonding behaviors depend on chain stiffness . . . . .	82
5.6	Summary and conclusions . . . . .	89
<b>6</b>	<b>Conclusions</b>	<b>91</b>
6.1	Summary of results . . . . .	91
6.2	Impact and applications . . . . .	92
6.3	Future work . . . . .	97
6.3.1	Experimental exploration of simulation results . . . . .	97
6.3.2	Simulation of additional polymerization schemes . . . . .	99
6.4	Closing remarks . . . . .	105

<b>A</b>	<b>Chapter 3 Supplemental Details and Results</b>	<b>106</b>
A.1	Simulation details . . . . .	106
A.1.1	Activation energy . . . . .	107
A.2	Polymerization kinetics and chain length distribution. . . . .	109
A.3	Markovian block length distribution . . . . .	110
A.4	Supplemental results for stiff chains . . . . .	110
A.5	Selective solvent results . . . . .	110
<b>B</b>	<b>Chapter 4 Supplemental Details and Results</b>	<b>116</b>
B.1	Simulation details . . . . .	116
B.1.1	Attractive interaction strengths and Lorentz-Berthelot combining . . . . .	117
B.1.2	Activation energies . . . . .	117
B.2	Calculation of the first coordination number from $g(r)$ . . . . .	119
B.3	Quantifying deviations from Markovian block statistics . . . . .	120
B.4	Local nematic order parameter for flexible chains . . . . .	120
B.5	Polymerized fraction . . . . .	123
B.6	Supplemental kinetics and length distributions . . . . .	124
<b>C</b>	<b>Chapter 5 Supplemental Details and Results</b>	<b>127</b>
C.1	Simulation details . . . . .	127
C.1.1	Langevin dynamics and viscosity . . . . .	127
C.1.2	Activation energies . . . . .	127
C.2	Chain length dependence of bonding statistics . . . . .	128
C.3	Deviations from Markovian predictions develop in time . . . . .	131
C.4	Clustering analysis . . . . .	131
C.5	System parameters and characteristic timescale . . . . .	133
C.6	Characteristic length in homopolymer system . . . . .	137
	<b>References</b>	<b>141</b>

# Chapter 1

## Introduction

### 1.1 Synthetic copolymers and sequence control

Since the landmark discovery of Staudinger<sup>1,2</sup> positing the existence of polymeric macromolecules, the study of polymers has seen enormous attention and tremendous advancement. More than a century of polymer science has led to the widespread adoption of polymeric materials, first as synthetic polymer rubbers and thermoplastics replaced organic rubbers or other materials in an industrial context, then as more tailored applications of properties unique to advanced polymer materials.<sup>3-5</sup> A particularly notable development was the introduction of copolymers, polymers composed of multiple monomer species.<sup>5</sup> Block copolymers, copolymers in which each comonomer species is isolated into repeated “block” regions along the chain, have seen a wealth of application in fields ranging from microelectronics manufacture<sup>6-8</sup> to targeted drug delivery<sup>9-11</sup> thanks to their propensity to self assemble into ordered structures on the nanoscale.<sup>8,12</sup> Block copolymers, however, require specific synthetic strategies to generate, as simultaneous polymerization of a mixture of comonomers results in “random”, or more properly, “statistical” copolymers, in which comonomer sequence and composition varies in a fashion related to known statistical laws and dependent on characteristics of the reaction kinetics and conditions.<sup>5,13,14</sup> It is the behavior of such polymerizations which we now explore.

Nature has evolved highly-tuned, molecular machinery to control the sequence of biological copolymers, from proteins to nucleic acids, and well-established methodologies exist for the controlled synthesis of such molecules.<sup>15</sup> By contrast, developing general synthetic methodologies for sequence regulation and sequence control in non-biological

copolymers remains an open challenge.<sup>14,16–19</sup> Throughout the last few decades, significant effort has gone into developing sequence controlled polymers – synthetic copolymers in which the primary sequence of monomer types along the chain are intentionally biased or controlled.<sup>14,16,19</sup> These efforts have led to a number of notable approaches that control or bias sequence, such as iterative synthesis with suitable deprotection/protection cycles for specific functional groups,<sup>20–22</sup> controlled radical polymerizations (CRP) that tune comonomer reactivities,<sup>23–27</sup> or switchable catalysts.<sup>28</sup> Despite these efforts, developing general synthetic methodologies for sequence biasing remains an open challenge in polymer science, as current methods for sequence control in synthetic copolymers lack the high degree of specificity or, crucially, the ease of application found in their biological counterparts.<sup>14,16,17,19,29,30</sup>

## 1.2 Copolymer sequence influences material properties and behaviors

Sequence control in synthetic polymers opens up a host of potential applications, both as macromolecular carriers of information, providing a route for developments in non-biological molecular information storage and synthetic biology,<sup>14,31–33</sup> and as highly functional materials due to the direct dependence of material properties and macromolecular morphologies on the polymer sequence. The unique material properties and self-assembly behaviors of copolymers, crucial to their numerous applications, are intrinsically related to their comonomer sequences and sequence distributions, and sequence defined polymers show tremendous potential for advanced materials design.<sup>29,30</sup> Various studies have shown that mechanical and rheological properties,<sup>24,34</sup> thermal and ionic conductivity,<sup>35–38</sup> glass transition and crystallinity,<sup>39–41</sup> morphology and phase behavior,<sup>24,37,42–47</sup> coacervation and electrostatic interactions,<sup>48,49</sup> optoelectronic properties,<sup>50–52</sup> and even sound and vibrational damping<sup>53,54</sup> are influenced by the primary monomer sequence, as evidenced by comparisons between polymers with different sequences but the same overall fractional monomer content. Further work has demonstrated the potential for targeted additions of specific monomers to a polymer sequence in order to obtain specific properties of interest.<sup>36</sup> The ability to synthetically control or bias copolymer sequences and their supramolecular organization will enable the informed design of novel and tunable materials for numerous applications.

### 1.3 Sequence influencing effects in copolymerization

In spite of the growing understanding of the fundamental relationship between sequence and material behaviors, sequences in non-biological copolymers are often not well-characterized. This is due to the dispersity introduced by standard synthetic methods,<sup>55</sup> as well as to limited understanding of factors that determine sequence development throughout the reaction.<sup>14</sup>

Though there has been marked progress made in developing specialized synthetic methods for sequence controlled polymers, progress towards a generalized methodology is hampered by the complexities inherent in even simple one-pot synthetic approaches.<sup>56,57</sup> Due to these complexities, the factors influencing sequence during these reactions are still not fully understood, limiting the extent to which they might be exploited to regulate sequence in these more "traditional" polymerizations.<sup>14,16,58</sup> Some previous work has explored such sequence influencing effects in a simplified model of a solvated step-growth polymerization among two monomer types, demonstrating that the presence of even relatively weak non-bonding attractions between monomers can induce an emergent microphase separation as the reaction proceeds.<sup>58</sup> When this happens, the self-assembly of the growing oligomers can locally enrich the concentration of certain species at the reaction site, thereby altering reaction kinetics and copolymer sequence. The resulting non-standard reaction kinetics can be further altered by the alignment of nascent oligomers with longer persistence lengths.<sup>59</sup>

This type of local concentration enrichment has previously been reported in both experimental and computational observations of the "bootstrap" effect,<sup>57,60–65</sup> in which an increase in the local availability of one monomer type shifts the relative kinetics of comonomer pairs during a reaction. Analogous behavior occurs in polymerization induced self-assembly (PISA), a widely studied synthetic approach in which a controlled radical polymerization drives self-assembly among nascent oligomers consisting of a growing block of one type of monomer attached to a pre-formed oligomer of a different type.<sup>66–68</sup> In these PISA formulations, an emergent phase separation driven by the polymerization reaction itself can locally enrich monomer concentrations, and thereby speed reaction rates, while also yielding desirable supramolecular morphologies.<sup>66,69–72</sup> Non-bonding interactions are a primary driver of assembly in PISA formulations, making the collective behaviors of monomer and oligomer species of particular relevance to the reaction product in such cases. While polymerization induced self-assembly has been intensively studied,<sup>66</sup> those studies are typically restricted to the chain extension of one monomer species onto a seed chain of another species. Less is known about how assembly behaviors impact the development of sequence

when more than one monomer species may be incorporated, particularly in the context of a step-growth reaction mechanism.

Despite the importance of such effects in the mesoscale behavior of copolymerization reactions, direct theoretical calculations of reaction kinetics often make fundamental assumptions that run counter to these observations of self-assembly and collective behaviors, assumptions such as truly homogeneous solutions with monomer diffusivity and reactivities independent of chain formation.<sup>73–75</sup> This is further complicated by the difficulty in experimentally characterizing the sequence of a non-biological polymer:<sup>14,16,55,76</sup> copolymers are often characterized by fractional composition alone, and even when some degree of sequence determination can be made, the development of sequence throughout the reaction presents further challenge to characterization. Molecular simulations provide a powerful tool to more quantitatively explore behaviors that run counter to current theoretical models, by removing many of these simplifying assumptions, and directly observing sequence development at a molecular level.<sup>42,58,63,77–80</sup> Reactive molecular dynamics simulations<sup>81–85</sup> in particular can allow for both observing sequence development throughout a reaction, as well as taking into account the influence of the reaction environment at a system level, producing realistic dynamics by directly calculating the trajectories of reactants from intermolecular potentials. Employing these simulation strategies will improve our understanding of emergent collective behaviors in polymerization, and would allow us to recognize the factors which enhance or minimize their effects, leading to new and broadly applicable avenues with which to approach the problem of controlling synthetic copolymer sequence and morphology.

## 1.4 Dissertation overview

This dissertation utilizes a reactive, coarse-grained molecular simulation approach to explore the development of primary sequence during a simultaneous, two-component, step-growth copolymerization. A particular target of this exploration is the emergent collective behavior of monomers and oligomer chains during polymerization due to non-bonding interactions or environmental effects. Through the model and simulation approach employed, we seek to understand the complex interplay between reaction driven assembly, polymerization kinetics, and sequence development, with an eye towards exploiting facets of this interdependence to develop new synthetic strategies for sequence controlled polymeric materials.

In Chapter 2, we describe the details of our coarse-grained model, including descriptions of the basic reactant constituents of our simulations and all non-bonded and bonded interaction potentials between our reactants. Further we describe the initial configuration and time-evolution sampling of our model, and the logic which determines how the reaction proceeds. Additionally, we describe some metrics for simulation analysis that are commonly utilized throughout the subsequent Chapters.<sup>†</sup>

In Chapter 3, we develop the fundamental kinetic theory of our model system from first principles of classical polymerization kinetics. We connect this ideal kinetic model to our simulation approach, and utilize this connection to explore how the heterogeneity produced by polymerization driven ordering and demixing causes a breakdown in the idealized homogeneity assumptions of standard kinetic theories of step-growth polymerizations.

In Chapter 4, we examine the impact of asymmetry in non-bonded intermolecular interactions, such as would be introduced by the presence of a solvent selective for one of the reacting species. Such a solubility preference alters the collective organization and polymerization driven demixing of the reactants, causing assembly to occur at different stages for the respective comonomer species. This change to assembly alters, in-turn, the kinetic and sequence impacts we have identified, as well as the composition profiles of the resulting assembled structures.

In Chapter 5, we examine in detail the impact of chain stiffness and nematic alignment on kinetic and sequence development in our reaction, an effect hinted at in the earlier Chapters. We explore how nematic ordering emerges in time, and how this coincides with the development of a characteristic block length within the system. We further examine how this characteristic block length shifts in response to changes in system parameters which alter the characteristic timescales of reaction and diffusion in the system. We then explore how this characteristic length arises from length-dependent bonding behaviors in oligomer aggregates, and how the structure and formation of these aggregates alters chain and block length statistics in a manner that differs between chains of different stiffness.

In Chapter 6 we provide a summary of our results, connecting them to the broader context of work in the study of copolymeric materials, sequence control, and polymerization induced phase separation and assembly. We then propose future routes of study, including both experimental methods targetting the observation and exploitation of the phenomenon we observe in our simulations, as well as extensions of our simulation model to other copolymerization schemes. Finally, we conclude with some brief thoughts regarding the

---

<sup>†</sup> We note here that some simulation or analysis details unique to work in specific Chapters are provided in the Chapters themselves or their respective Appendices.

potential for further study and utilization of self-assembly and self-organization principles to advance the field of sequence controlled polymers and thereby aid in the design and broad application of advanced polymeric materials.

# Chapter 2

## Methods

### 2.1 Model details

#### 2.1.1 Monomer structure and interactions

To simulate an irreversible, step-growth polymerization, we make use of a coarse-grained, reactive representation of a step-growth copolymerization, a model previously developed in Ref. [58] and expanded in subsequent works.<sup>59,86,87</sup> It is based on common “bead and spring” models for polymers, which have been employed in the study of a wide variety of material properties and have yielded well-validated results.<sup>88,89</sup> A single bifunctional monomer is represented as a simplified “bead.” Each bead is composed of three particles: a central particle connected to two external particles via harmonic bonds. These harmonic bonds control the spatial and angular separation of the three particles in the bead, modeling the internal configuration of the monomer and determining the chain stiffness. The central particle of each monomer bead contains the majority of the mass and determines the monomer species, denoted either **A** or **B** to represent the two comonomers taking part in the copolymerization. The external particles represent the reactive functional groups and form bonds with the external particles of other monomers when the reaction conditions are met.

---

Portions of the model described in this Chapter were initially developed in Ref. [58], and have been further developed by the author and coworkers in Ref. [59, 86, 87].

Unbound monomers interact with one another via two potentials, each associated with one of the particle types within the monomer bead. The central, type **1** particle determines the primary non-bonding intermolecular interactions within the system, by means of a modified Lennard-Jones (LJ) potential.<sup>58</sup>

$$E_{\text{LJ}}(\mathbf{1}, \mathbf{1}') = \begin{cases} 4\epsilon_{\text{att}(\mathbf{1}, \mathbf{1}')} \left[ \left( \frac{\sigma}{r_{(\mathbf{1}, \mathbf{1}')}} \right)^{12} - \left( \frac{\sigma}{r_{(\mathbf{1}, \mathbf{1}')}} \right)^6 \right] & r_0 \leq r_{(\mathbf{1}, \mathbf{1}')} < 2.5\sigma \\ 4\epsilon_{\text{rep}} \left[ \left( \frac{\sigma}{r_{(\mathbf{1}, \mathbf{1}')}} \right)^{12} - \left( \frac{\sigma}{r_{(\mathbf{1}, \mathbf{1}')}} \right)^6 \right] + c & r_{(\mathbf{1}, \mathbf{1}')} < r_0, \end{cases} \quad (2.1)$$

where  $r_{(\mathbf{1}, \mathbf{1}')}$  is the distance between the two interacting monomer centers **1** and **1'**,  $r_0$  is the separation at the minimum of the potential ( $r_0 = 2^{\frac{1}{6}}\sigma$ ), and  $c$  is a constant which ensures continuity between the attractive and repulsive portions of the potential. The repulsive portion, controlled by the well depth  $\epsilon_{\text{rep}}$  is kept constant regardless of monomer identity, while the attractive portion changes in strength based on monomer identity, as governed by the well depth  $\epsilon_{\text{att}(\mathbf{1}, \mathbf{1}')}$ . This has the advantage of allowing attractive interactions to depend upon monomer identity – i.e., by setting  $\epsilon_{\text{att}(\mathbf{1}, \mathbf{1}')} = \epsilon_{\text{AA}}$ ,  $\epsilon_{\text{att}(\mathbf{1}, \mathbf{1}')} = \epsilon_{\text{AB}}$ ,  $\epsilon_{\text{att}(\mathbf{1}, \mathbf{1}')} = \epsilon_{\text{BB}}$  for **A** to **A**, **A** to **B**, and **B** to **B** monomer interactions, respectively – while simultaneously maintaining fixed steric repulsions for each species. The model can therefore explore cases with zero attractive interactions, while maintaining fixed monomer size-exclusion afforded by the repulsive interactions of  $\epsilon_{\text{rep}}$ . For notational simplicity, we denote the attractive interaction strengths between monomers of species  $i$  and  $j$  as  $\epsilon_{ij}$  throughout this work.

The type **2** interaction particles of each monomer bead represent the reactive functional groups of the monomer. In order to model a reactive process with a tunable activation energy, we define an additional, soft and very short-ranged repulsive potential between type-**2** particles via:

$$E_{\text{soft}}^{2,2'} = \begin{cases} \frac{E_{\text{barr}}^{ij}}{2} \cos \frac{\pi(r_{2,2'} - d_{\text{bond}})}{d_{\text{on}} - d_{\text{bond}}} + \frac{E_{\text{barr}}^{ij}}{2} & d_{\text{bond}} < r_{2,2'} < d_{\text{on}} \\ 0 & r_{2,2'} \geq d_{\text{on}}, \end{cases} \quad (2.2)$$

in which  $d_{\text{on}}$  is the cutoff distance for the potential and  $E_{\text{barr}}^{ij}$  is the height of this contribution to the reaction barrier, which can be adjusted across different  $ij$  monomer pairs. Additional details on the use of this potential in modeling reaction activation energy is provided in Sec. 2.2 below.

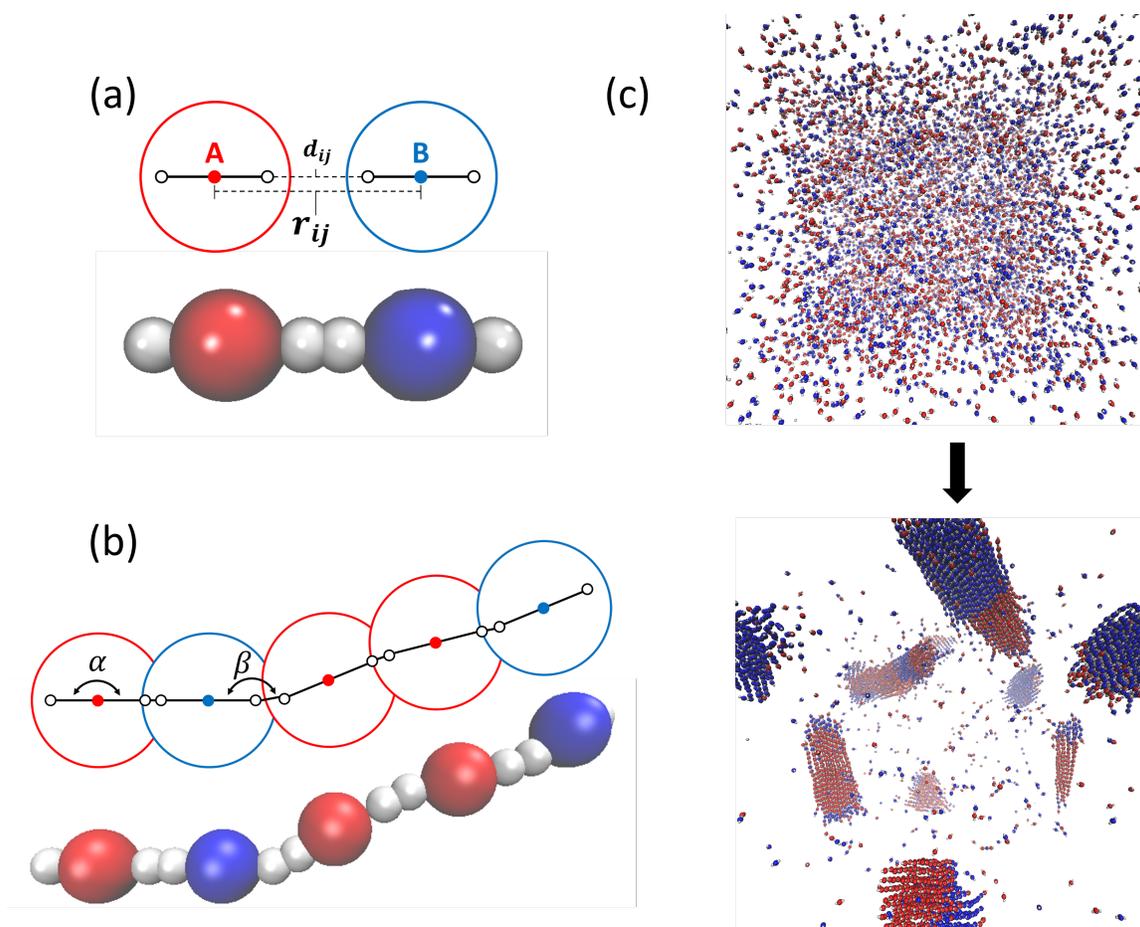


Figure 2.1: **Coarse-grained model schematic and reaction visualization.** (a) A schematic and associated visualization of each monomer type in the copolymerization simulation. Monomers are represented as coarse-grained spherical beads containing three particles as shown. Central particles define the monomer type, either A or B, while the external particles represent reactive moieties. Both the center to center distance  $r_{ij}$ , which controls the intermolecular interactions via the LJ potential, and the interaction particle distance  $d_{ij}$ , which determines the activation energy and bonding events, are shown. When  $d_{ij} \leq 0.2\sigma$ , an irreversible bond is formed. (b) A growing oligomer chain. Within the schematic, the intra-monomer angle  $\alpha$  and the inter-monomer angle  $\beta$  are shown, each with associated angular harmonic potentials. The strength of the  $\alpha$  angular potential determines the stiffness of the chain. (c) A visualization of the initial and final system states for a single step-growth copolymerization simulation.

### 2.1.2 Symmetric and asymmetric effective inter-monomer interactions

For the majority of simulations performed in this work we model what we term “symmetric” attractive interactions, where attractions between like monomer species are equivalent and  $\epsilon_{AA} = \epsilon_{BB} \equiv \epsilon_{AA,BB}$ . With this symmetric constraint, we consider situations where all monomer attractions are equivalent, *i.e.*,  $\epsilon_{AA,BB} = \epsilon_{AB} \equiv \epsilon_{\text{all}}$ , and those in which attractions between like monomers are different than those between unlike monomers, *i.e.*,  $\epsilon_{AA,BB} \neq \epsilon_{AB}$ . In Chapters 3&5 we employ such symmetric attractive interactions, and examine emergent kinetic and phase behaviors of the system in such cases. The precise values of  $\epsilon_{AA,BB}$  and  $\epsilon_{AB}$  used are described in these Chapters (Ch. 3&5) and their associated Appendices (Appx. A&C).

In Chapter 4, we explore copolymerizations in which monomers have asymmetric non-bonded interaction strengths, by imposing the condition  $\epsilon_{AA} > \epsilon_{BB}$ . This choice of attractive interaction strengths provides a stronger driving force for **A** monomers to self-associate than for **B** monomers – although we note that even at the maximum attraction strengths investigated here, all monomers remain well dissolved and do not begin to aggregate until a significant number of nascent oligomers form. This alteration to the attractive interactions for each monomer species is akin to that produced by a difference in solvent affinity, wherein one monomer, more solvophobic than the other, has a greater propensity for self-association. As solvent molecules are not explicitly represented in our system, the relative solvent affinities of the monomers are captured through these asymmetric LJ interactions between each monomer pair: the stronger like-monomer attractive interactions of species **A** make it the more solvophobic species as compared to species **B**, where weaker self-attractions produce a more solvophilic character. In order to determine the interactions between **A** and **B**, we define  $\epsilon_{AB} = \sqrt{\epsilon_{AA}\epsilon_{BB}}$ , making use of standard Lorentz-Berthelot combining rules for LJ interactions (which is equivalent to the Kong combining rules for our system).<sup>90,91</sup> With these conditions, a range of attraction strengths were chosen such that the maximum  $\epsilon_{AA}$  value was set to  $1k_B T$ , where signs of self-assembly and sequence altering effects were observed in prior work,<sup>58</sup> while the maximum  $\epsilon_{BB}$  value remained much lower at  $0.5k_B T$ . The chosen values for  $\epsilon_{AA}$  and  $\epsilon_{BB}$ , and the associated Lorentz-Berthelot values for  $\epsilon_{AB}$ , are presented in Table 2.1. For each set of attraction strengths explored, persistence lengths of  $l_p = 3.5$  and  $l_p = 16.5$  were tested, corresponding to flexible and stiff polymer chains as controlled by the angular harmonic potential. A detailed discussion of the persistence length calculation is available in prior work.<sup>59</sup>

$\epsilon_{AA} (k_B T)$	$\epsilon_{AB} (k_B T)$	$\epsilon_{BB} (k_B T)$
1.0	0.32	0.1
1.0	0.55	0.3
1.0	0.71	0.5
0.75	0.27	0.1
0.75	0.47	0.3
0.75	0.61	0.5
0.5	0.22	0.1
0.5	0.39	0.3
0.5	0.5	0.5

Table 2.1: **Lennard-Jones attractions.** Columns show  $\epsilon_{AA}$ ,  $\epsilon_{AB}$ , and  $\epsilon_{BB}$ , which are the well-depths for the attractive portion of the LJ interactions (see Eq. 1 in the main text) for **A** to **A**, **A** to **B** and **B** to **B** monomers, respectively. Values for  $\epsilon_{AA}$  and  $\epsilon_{BB}$  were chosen to reflect a copolymerization in which monomer species **A** is more solvophobic than species **B**. From the values of  $\epsilon_{AA}$  and  $\epsilon_{BB}$ ,  $\epsilon_{AB}$  values were calculated according to Lorentz-Berthelot rules.<sup>92</sup> For all combination of attraction strengths, the strength of repulsive interactions,  $\epsilon_{rep}$ , is kept at a constant value of  $1.33 k_B T$ , irrespective of monomer identity.

### 2.1.3 Radial and angular harmonic potentials describe bonded interactions

Bonded interactions in the model, the “springs” of our bead and spring model, are described through the use of harmonic potentials. The harmonic functional form of these potentials describes all bonds in the system, both within monomer beads and between two bonded monomers once a reaction occurs, with differences in the parameters used distinguishing the intramonomer and intermonomer bond types. Two harmonic bonds control the interparticle spatial separation and angular structure of the particles, which together model the internal configuration of the monomer. Adjusting the strength of the angular harmonic potential alters the rigidity of the monomer and can be used to control the persistence length of the associated polymer.

Intramolecular and intermolecular angles are governed via the harmonic potential:

$$E_{\text{angle}}(\theta_{ijk}) = K_{ijk}^{\text{angle}}(\theta_{ijk} - \theta_0)^2 \quad (2.3)$$

in which  $\theta_{ijk}$  is the angle between particles  $i$ ,  $j$  and  $k$ ,  $\theta_0$  is the equilibrium angle, and  $K_{ijk}^{\text{angle}}$  is the spring constant for the angle. For the intramonomer angle between particle **2-1-2'**, angle  $\alpha$  in Fig. 2.1b,  $\theta_0 = 180^\circ$  and the spring constant is set to  $K_{212'}^{\text{angle}} = 5 \text{ } \epsilon \text{ rad}^{-2}$  and  $K_{212'}^{\text{angle}} = 50 \text{ } \epsilon \text{ rad}^{-2}$  for the flexible chains and stiff chains, respectively. For the intermonomer angle between particles **1-2-2'**, angle  $\beta$  in Fig. 2.1b,  $\theta_0 = 180^\circ$  and  $K_{212'}^{\text{angle}} = 100 \text{ } \epsilon \text{ rad}^{-2}$  for both flexible chains and stiff chains.

Similarly, both intramonomer and intermonomer distances are governed via the harmonic potential:

$$E_{\text{bond}}(r_{ij}) = K_{ij}^{\text{bond}}(r_{ij} - r_0)^2, \quad (2.4)$$

in which  $r_{ij}$  is the distance between particles  $i$  and  $j$ ,  $r_0$  is the equilibrium distance, and  $K_{ij}^{\text{bond}}$  is the spring constant for the bond. For the intramonomer bond between particles **1** and **2**,  $r_0 = 0.4\sigma$ , and the spring constant is set to  $K_{12}^{\text{bond}} = 2000 \text{ } \epsilon \text{ } \sigma^{-2}$ . For the intermonomer bond between particles **2** and **2'**,  $r_0 = 0.15\sigma$ , and the spring constant is set to  $K_{22'}^{\text{bond}} = 6000 \text{ } \epsilon \text{ } \sigma^{-2}$ .

## 2.2 Simulation progression and bond formation

### 2.2.1 Sampling via Langevin dynamics

Simulations begin with equal numbers of **A** and **B** monomers distributed randomly throughout the simulation volume. Time evolution occurs according to Langevin dynamics.<sup>93</sup> Sampling the system via the Langevin equation not only produces realistic dynamics from the potentials in the system, but also incorporates a viscous drag force and random collision forces characteristic of the solution-state and serves as a thermostat for the system. Adjusting the parameters of the Langevin equation allows us to reproduce solution-state dynamics for a range of different solution viscosities (see Appendix C Sec. C.1.1 for details). The system is first allowed to equilibrate with only non-bonded interactions before bonding is turned on. It is important to note that, for all simulation parameters explored in this work, monomers remain well-dispersed in the solution phase prior to polymerization, only beginning to aggregate as the reaction proceeds and oligomers lengthen due to the change to the associated free energy of mixing of longer chains, as described by Flory-Huggins theory.<sup>73</sup> After the equilibration period, attractive interactions and bond formation are switched on, and the system is progressed to a reaction extent of  $p = 0.9$ , i.e., to the point where 90% of bond formation has occurred. Fig. 2.1c shows the initial and final system state for a representative simulation under these conditions.

### 2.2.2 Modeling a step-growth reaction mechanism

Reaction events are modeled by defining a specific cutoff distance  $d_{\text{bond}}$ . When the external particles of two monomers cross within this distance, overcoming both the fixed steric repulsions from the central particles and the additional repulsions between the external reactive particles, an irreversible harmonic bond is formed (Fig. 2.1a). Any two unreacted functional groups may form a bond with each other and each functional group is capable of forming one such new bond, and the reaction thus proceeds through a step-growth mechanism to produce unbranched, linear polymers. Similar reaction cutoff approaches have been used in a variety of polymerization simulations, and have been shown in coarse grained models to successfully capture key reaction characteristics from fully atomistic molecular simulations.<sup>81,82,94</sup>

The potential defined in Eq. 2.2 is designed to serve as an additional, adjustable reaction barrier to bond formation. When any two type **2** particles overcome this barrier, as well as

the repulsive LJ interactions between their respective monomer centers, and come within a defined distance  $d_{\text{bond}}$ , an irreversible bond is formed. Upon bond formation, all non-bonded interactions between directly bound monomers are switched off. The total activation energy for a reaction,  $E_a^{ij}$ , is then composed of contributions from the non-bonded interactions between the reactants – which depend upon the monomer identities  $i$  and  $j$  of the two reacting species – and from geometric constraints arising from the intramonomer harmonic bonds of the respective reacting monomers. Specifically, both the steric repulsions between **1** - **1'** particles from Eq. 2.1 (the strength of which is fixed via  $\epsilon_{\text{rep}}$  and is held constant across all simulations) and the very short-range repulsions between **2** - **2'** particles from Eq. 2.2 (which we vary) contribute to the overall magnitude of the activation energy for polymer bond formation. If this energetic barrier is overcome, and the distance between **2** - **2'** is less than  $d_{\text{bond}}$ , the bonding distance, an irreversible bond will be formed as governed via a harmonic potential. Due to the irreversible nature of the polymerization modeled, bond breakage does not occur and simulations are inherently non-equilibrium, representing the behavior at the limit of strong bond formation. Unless otherwise noted, the attractions and activation energies of monomer **A** and **B** throughout this work are set to be symmetric so that  $\epsilon_{\text{AA,BB}} = \epsilon_{\text{AA}} = \epsilon_{\text{BB}}$  and  $E_a^{\text{AA,BB}} = E_a^{\text{AA}} = E_a^{\text{BB}}$ . As discussed in Sec. 2.1 above, the  $\epsilon_{\text{AA,BB}}$  and  $\epsilon_{\text{AB}}$  that are varied refer to the strength of the LJ attractions,  $\epsilon_{\text{att}}$ , while the LJ repulsion term, specified by  $\epsilon_{\text{rep}}$ , is held constant in all cases to maintain appropriate steric repulsion.

### 2.2.3 Reaction activation energies

Throughout this work we make use of the concept of activation energy,<sup>95</sup> which we define here as the maximum total potential energy experienced by two monomers during bond formation. This definition allows us to utilize the Arrhenius formalism in kinetic analyses of our system, as discussed in Ch. 3, while also leveraging the advantages of a particle-based, dynamical model to measure the activation energy directly observed in our simulations in a straightforward manner. Activation energies are therefore calculated from the total potential energy of the two bonding monomers arising from the LJ potential between type **1** particles, the soft repulsive potential between type **2** particles, and the intramonomer bond lengths and angles. The values of each of these contributions above their minimum energy are summed at the time of bond formation to obtain the total activation energy for the bonding event. This calculation can be repeated for a large ensemble of bonding events across multiple simulations for a given set of reaction conditions, providing an effective measure

of the ensemble averaged activation energy.

In previous work,<sup>58,59</sup> and in Ch. 4&5, we maintained a fixed activation energy across different monomer species combinations in order to isolate and investigate the effects of relative non-bonded attractions on polymerization. In Chapter 3, we seek to understand the combined effects of non-bonded attractions and relative activation energy on copolymerization by investigating the kinetic and sequence behaviors of our model over a range of non-bonded attractions and activation energies. Therefore, we vary  $\epsilon_{\text{att}}$  and  $E_a$  values to explore the systems at different attraction strengths over a range of  $E_a$  values. To vary the activation energy in this study, only the additional repulsions between type-2 particles are varied (by adjusting the value of  $E_{\text{barr}}^{ij}$  in Eq. 2.2), while the other two components are kept fixed. Additional details on the measurement of activation energy for each of the conditions explored in these Chapters are available in the Appendices for the respective Chapters. In summary, we find that adjusting the strength of the repulsive potential defined in Eq. 2.2 successfully alters the total activation energy experienced during a bond formation in a species dependent manner. Further, we confirm that the activation energy is not affected by other factors varied in our simulations in Ch. 4&5: that changes to attractive interaction strengths do not impact the activation energy for different monomer species combinations, and that activation energies are consistent for oligomer chains of different stiffnesses. That the activation energy behaves in the manner expected allows us to distinguish the impact these variations have: altering the structure of the system and reactants in the case of attractive interactions and chain stiffness variation, or changing the reaction pathway itself in the case of a change to the activation energy directly.

## 2.3 System descriptors and simulation analysis metrics

### 2.3.1 Chain length and block length statistics

In characterizing the results of our simulation, two critical metrics are the chain length and block length distributions. These probability distributions describe, respectively, the probability of observing a chain of a given length within the system, or a block of repeated monomers of the same type of some length within the oligomer chain in the system. One of the key advantages of our model is that it allows us to collect these distributions from direct observation of simulation results: simply counting the length of chains and blocks in the system over an ensemble of simulations and generating a histogram directly. This allows

us to compare our observations to commonly used descriptive statistical theories of polymerization, to better understand conditions under which the assumptions underlying these theories begin to break down.<sup>73</sup> Here we describe the theoretical underpinnings of such statistical descriptions, describing the concepts behind the Flory-Schulz theory for chain length distribution, and deriving an associated prediction for the block length distribution among the oligomers based on a Markovian statistical model.

We begin with the probability distribution for chain lengths in a step-growth polymerization, which is given by the Flory-Schulz distribution.<sup>73</sup> The Flory-Schulz distribution is a well-studied and validated distribution which describes the probability,  $P(x)$  of obtaining chains of length  $x$  in a step-growth polymerization reaction in terms of the reaction extent,  $p$ , as:

$$P(x) = (1 - p)p^{x-1}. \quad (2.5)$$

Distributions calculated from Eq. 2.5 are shown throughout the text for comparison to the observed distributions of oligomer chain lengths.

Under Flory's assumption of equal reactivity, the probability of having a like or unlike bond is only based on the identities of the two bonding monomers, as reactivity does not depend on the length or properties of a molecular chain attached to either of the two reacting monomers. This dependence of reactivity and bond formation on only the identities of the most recent monomer and next monomer to be added means that the sequence can be treated as a Markov chain. To obtain an expression for the probability distribution of contiguous blocks of either **A** or **B**, we define analogous equations to Equation 2.5, but specific to **A** or **B**. Just as  $p$  in Eq. 2.5 is the probability that a given monomer has reacted with any other given monomer, we now define  $p_A$  as the probability that a given monomer of **A** has reacted with another **A** monomer. Similarly,  $p_B$  the probability that a given monomer of **B** has reacted with another **B** monomer. Then, for the distributions of **A** blocks of length  $n_A$  and **B** blocks of length  $n_B$ , we obtain:

$$\begin{aligned} P(n_A) &= (1 - p_A)p_A^{n_A-1}; \\ P(n_B) &= (1 - p_B)p_B^{n_B-1}. \end{aligned} \quad (2.6)$$

The resulting Markovian expression, Eq. 2.6, is used as a point of comparison for our observed sequence behavior throughout this work, and holds regardless of the symmetry of interactions between **A** and **B**. In the case of symmetric interactions, we can additionally

consider the probability of observing a repeat block of length  $n$ , independent of the type of block. In this case, we may use symmetry to calculate  $p_A = p_B = pp_{AA,BB}$ , where  $p_{AA,BB}$  is the frequency of nearest neighbor pairs of **AA** or **BB** for sequences in the system. This results in the expression in Eq. 2.6 being equivalently expressed as:

$$P(n_A) = (1 - pp_{AA,BB})pp_{AA,BB}^{n-1}. \quad (2.7)$$

From the Flory-Schulz distribution it is possible<sup>5,73</sup> to derive a corresponding expression for the system dispersity,  $\mathcal{D}$ , for a given reaction extent  $p$ , which also provides an important comparison for the simulation results. We start with  $\mathcal{D}$  defined as:

$$\mathcal{D} = \frac{M_w}{M_n}, \quad (2.8)$$

which is the ratio of the weight-averaged molar mass,  $M_w$ , to the number-averaged molar mass,  $M_n$ . These are each obtainable from the Flory-Schulz distribution,  $P(n)$  in Eq. 2.5, using the definitions:<sup>5,73</sup>

$$M_n = M_0 \sum_n nP(n); \quad (2.9)$$

$$M_w = \sum_n w_n M_n = M_0 \sum_n w_n nP(n). \quad (2.10)$$

Here  $w_n$  is the weight-fraction of  $n$ -mers given by:

$$w_n = \frac{(nM_0)N_n}{M_0N_0} = \frac{nN_n}{N_0}, \quad (2.11)$$

for  $N_n$  total  $n$ -mers and  $N_0$  initial monomers with a molar mass of  $M_0$ . From these definitions the following identities are used:

$$\sum_x xp^{x-1} = (1-p)^{-2} \text{ for } p < 1; \quad (2.12)$$

$$\sum_x x^2 p^{x-1} = (1+p)(1-p)^{-3} \text{ for } p < 1. \quad (2.13)$$

Using the expressions in Eq. 2.12 and 2.13, the mass-averaged and weight-averaged molar mass expressions reduce to:

$$M_n = M_0 \frac{1}{1-p}; \quad (2.14)$$

$$M_w = M_0 \frac{1+p}{1-p}. \quad (2.15)$$

Finally, combining Eq. 2.14&2.15 with the definition of  $\mathcal{D}$  in Eq. 2.8 we obtain the expression:

$$\mathcal{D} = 1 + p. \quad (2.16)$$

The resulting expression<sup>5,73</sup> in Eq. 2.16 is utilized throughout the text and is compared to the observed dispersities in our simulations.

For the derivation of Eq. 2.16 we made use of the Flory-Schulz distribution for the entire chain given in 2.5. To obtain an expression for the expected dispersity in block lengths, we replace the Flory-Schulz distribution for chain lengths with the expected distribution of block lengths using a Markovian model, namely  $P(n_A)$  or  $P(n_B)$  as defined in Eq. 2.6. This leads to the following expressions for block length dispersity:

$$\begin{aligned} \mathcal{D}_A &= 1 + p_A; \\ \mathcal{D}_B &= 1 + p_B. \end{aligned} \quad (2.17)$$

### 2.3.2 Quantifying deviations from Flory-Schulz and Markovian predictions

To quantitatively explore the deviation from Flory-Schulz and Markov statistics in our system, we compared the predicted chain and block length distribution with the results of our simulations. We calculated the Wasserstein metric,<sup>96,97</sup>  $W_1$ , as a measure of the statistical distance between the Markovian distributions and our observed block length distributions. For two discrete, univariate probability distributions  $p(x)$  and  $q(x)$  with cumulative distribution functions  $P(x)$  and  $Q(x)$ , the first-order Wasserstein distance,  $W_1$  may be defined as:

$$W_1(p(x), q(x)) = \int_{-\infty}^{\infty} |P(x) - Q(x)| dx. \quad (2.18)$$

### 2.3.3 Chain length dependence of bonding statistics

The fundamental simplifying assumption at the heart of the Flory-Schulz expression in Eq. 2.5 is the so-called ‘‘Flory’s equal reactivity principle’’ which supposes that the likelihood of reaction between two polymerizing species is completely independent of the chain

length of either reactant. Though Flory himself acknowledged that this principle would breakdown under heterogeneous reaction conditions,<sup>73</sup> such as those caused by the emergent demixing in the system studied in this work, this assumption serves as a worthwhile starting point for the purposes of highlighting the chain-length dependence in bonding which we observe.

In order to translate this assumption into probabilities of particular chain length pairs forming a bond at each stage of the reaction, allowing us to thereby compare our simulation results to those expected from the equal reactivity case, we performed a series of simple Monte Carlo simulations. Each simulation begins with a collection of 7200 monomers, just as in the full Langevin dynamics simulation of our system. We then choose two monomers, with uniform probability, and cause them to react. This process is performed iteratively, maintaining a uniform selection probability for all reactants, until a reaction extent of  $p = 0.9$  is reached. This was repeated over a series of 250 Monte Carlo simulation trials. The resulting bonding events were tabulated and split into three stages of the reaction  $p \leq 0.3$ ,  $0.3 < p \leq 0.6$ , and  $0.6 < p \leq 0.9$ , corresponding to the early, middle, and late stages of the reaction respectively. The histogram of the resulting bonding events is provided and discussed in Fig. 5.4a in Chapter 5. The chain length distribution we obtain from this Monte Carlo simulation process matches the Flory-Schulz prediction from Eq. 2.5 nearly identically, as shown in Fig. 2.2 below. We note here that this precise matching of Flory-Schulz predictions is also seen in Langevin dynamics simulations of our full model under conditions with sufficiently weak non-bonded attractions, as reported in our previous works.<sup>58,86,87</sup>

### 2.3.4 Quantifying chain stiffness and nematic alignment

As a quantitative measure of the chain stiffness, we consider the persistence length  $l_p$  defined as:

$$\langle \cos \theta \rangle = \exp(-l/l_p), \quad (2.19)$$

where  $l$  is the contour length distance between two points on the polymer chain, and  $\theta$  is the angle between the tangent lines drawn at each of these two points. An average is taken over an equilibrium ensemble of configurations at each contour length distance of  $l$ . All simulations in this work were run with an angular potential chosen to produce persistence lengths of  $l_p = 3.5$  monomer units for “flexible” chains or  $l_p = 16.5$  monomer units for “stiff” chains. These persistence lengths correspond to the lowest and highest values of  $l_p$

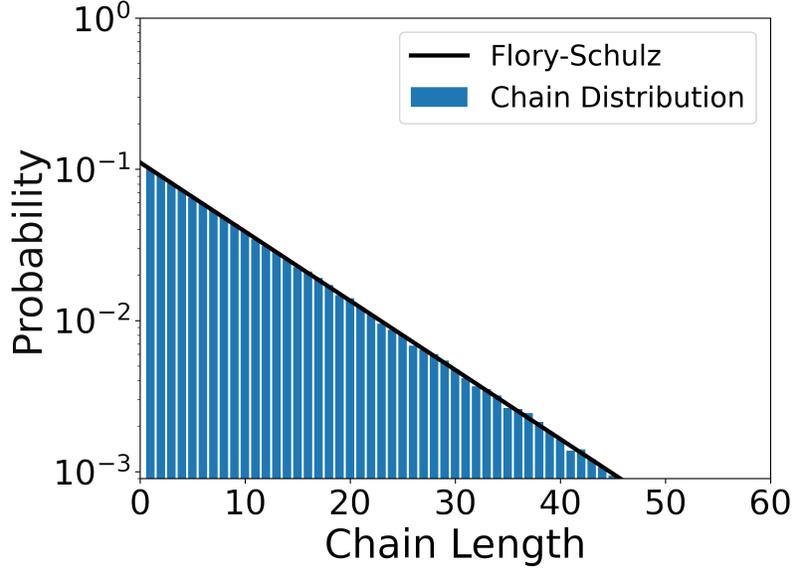


Figure 2.2: **Chain length distribution from Monte Carlo simulation of equal reactivity.** The observed chain length distribution at  $p = 0.9$  for 250 Monte Carlo simulation trials with uniform reaction probability. The predicted distribution from Flory-Schulz theory is also shown.

explored in our previous work.<sup>59</sup>

Persistence length determination was previously presented in Ref. [59]. Single chains of purely repulsive 100mers at several chain stiffnesses were simulated in order to calculate the persistence length,  $l_p$ , that corresponds to each value of  $K_{212}^{\text{angle}}$ . The value of  $l_p$  in each case was calculated using the relation  $\langle \cos \theta \rangle = \exp(-l/l_p)$ . Here,  $l$  is the contour length distance between two points on the polymer chain, and  $\theta$  is the angle between the tangent lines drawn at each of these two points. Persistence lengths are tabulated in monomer units,  $l = \Delta n$ , where  $n$  is simply the ordinal number of each monomer in the chain. An average is taken over an equilibrium ensemble of configurations at each contour length distance of  $l$ .

In order to characterize the local alignment of oligomer chains, we define a local nematic ordering parameter  $\overline{S_{\text{local}}}$  as shown:

$$\overline{S_{\text{local}}} = \left\langle \frac{3 \cos^2 \theta - 1}{2} \cdot \mathbf{1}_{\text{aggregates}} \right\rangle. \quad (2.20)$$

Here  $\theta$  is the angle between the orientation of a single monomer within the aggregate

---

and its local director, a vector taken as the average of all monomer orientations within the aggregate. The ensemble average, over each monomer in the system, is taken of the second order Legendre polynomial  $P_2(\cos(\theta))$  multiplied by an indicator function  $\mathbf{1}_{\text{aggregates}}$ . The function  $\mathbf{1}_{\text{aggregates}}$  takes the value 1 for monomers in aggregates and 0 elsewhere. For this calculation a monomer is considered to be within an aggregate if there are twelve or more neighboring monomers within a distance of  $2.5\sigma$ .

## Chapter 3

# Collective Reactant Behaviors and Kinetic Impacts

### 3.1 Introduction

Within a simple, one-pot copolymerization, the key sequence-influencing factors can be captured by the various rate constants of polymer bond formation between the monomer species, provided that Flory's principle of equal reactivity holds.<sup>73</sup> In this case, as the reaction proceeds, rate constants remain unchanged and the reacting species remain heterogeneously distributed. The dominant factor in these rate constants is the activation energy for bond formation between two specific monomers, and its influence can be modelled using the Arrhenius equation. Mayo-Lewis theory states that ratios of these rate constants, which are referred to as reactivity ratios, govern how different monomers are statistically incorporated into the copolymer chains.<sup>74</sup> For instance, in a system composed of **A** and **B** monomers, reactions between **A-A**, **B-B**, **A-B**, and **B-A** are governed by their respective rate constants,  $k_{AA}$ ,  $k_{BB}$ ,  $k_{AB}$  and  $k_{BA}$ . Their reactivity ratios,  $r_A$  and  $r_B$ , are defined as  $r_A = (k_{AA}/k_{AB})$  and  $r_B = (k_{BB}/k_{BA})$ . The product  $r_A r_B$  can then be used to predict sequence:  $r_A r_B \ll 1$  yields alternating copolymers;  $r_A r_B \gg 1$  yields block copolymers; and  $r_A r_B = 1$  yields a fully random copolymer. Reactivity ratios can even be used to

---

This Chapter has been adapted with minor alterations, with permission, from Ref. [87]. Copyright 2022 American Chemical Society. Ryan L. Hamblin contributed to the investigation, software, formal analysis, visualization, and writing. Nhu Q. Nguyen contributed to the investigation, software, formal analysis, visualization, and writing. Kateri H. DuBay contributed to the conceptualization, methodology, supervision, formal analysis, and writing.

predict more complex sequence statistics, such as the composition drift that occurs in a chain-growth mechanism when one species is consumed more rapidly than another.<sup>98–100</sup>

However, under certain conditions, ratios of static reaction constants are not sufficient to describe the polymerization kinetics that determine sequence. As Flory recognized, reaction kinetics become more complicated when there are inhomogeneities in the concentrations of the reacting species.<sup>73</sup> Under these circumstances, reaction rates cannot be modelled using static reaction constants and bulk concentration. Such conditions occur in the bootstrap effect<sup>60,61,65,101</sup> and in polymerization induced self-assembly (PISA).<sup>66,102,103</sup> In these cases, the local chemical composition at the reaction site differs from bulk concentrations, thereby altering the kinetics of bond formation.

Prior work on step-growth copolymerizations has demonstrated that relatively low effective intermonomer attractions gave rise to concentration heterogeneities as the nascent chains lengthen,<sup>58</sup> a trend predicted by Flory-Huggins theory.<sup>104,105</sup> These nascent chain assemblies led to an increase in the polymerization rate and a shift in the dispersity away from the ideal behaviors predicted by Flory.<sup>73</sup> These results demonstrated that even small variations ( $\leq 1k_B T$ ) in the effective inter-monomer attractions can influence sequence statistics in a way that cannot be accounted for by static reactivity ratios.

While these previous findings suggest a promising route to bias copolymer sequence within a one-pot step-growth reaction, activation energies are rarely the same across different combinations of comonomers. In this Chapter, we investigate the combined effects of differing non-bonded attractions and differing activation energies on sequence and self-assembly within our previously developed model of an irreversible **A,B** step-growth copolymerization.

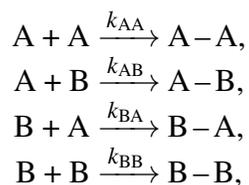
We first detail the kinetic behavior of our system under the "ideal" conditions of Flory<sup>73</sup> and Carothers,<sup>75</sup> and work out the dependence of the effective rate constants within the model on both the non-bonded attractions and the activation energies. In Section 3.3, we relate these effective rate constants to the expected sequence statistics and investigate how the sequence statistics from our copolymerizations compare for cases with varying attractions and activation energies. In Section 3.4, we demonstrate evidence of sequence biasing that persists beyond nearest neighbors, showcasing how the emergent reorganization of reactants leads to long-range sequence effects. In Section 3.5, we examine chain length and block length statistics for our system, describing how the combined effect of differences in inter-monomer attractions and activation energies lead to marked deviations from Flory-Schulz and Markov chain descriptions. Finally, in Section 3.6, we briefly summarize the results of this Chapter, and discuss how these emergent phase behaviors could be utilized

to bias comonomer sequence.

### 3.2 Ideal second-order kinetics of step-growth copolymerization

We begin by examining the kinetics of our model reaction in the interaction regime of  $\epsilon_{\text{att}} \leq 0.25k_{\text{B}}T$ , where we have previously observed ideal step-growth kinetics, such that the equal reactivity principle holds throughout the polymerization and the reaction kinetics can thus be modeled via a static rate constant and bulk reactant concentrations.<sup>58</sup> Specifically, we employ the Flory-Carothers<sup>73,75</sup> and Arrhenius<sup>95</sup> equations to work out, under these ideal conditions, the dependence of our simulated reaction rates on the non-bonded attractions and activation energies in the model.

$X_n$  **should increase linearly with time, with a slope that depends on the model's effective rate constant.** In the copolymerization of monomers **A** and **B**, reactions occur between the unreacted functional groups, denoted A and B respectively. We consider the case in which each of the reactive moieties of our bifunctional monomer species are equivalent, that is both reactive end groups of a given monomer **A** are denoted A and react equivalently, and similarly for end groups of monomer **B** denoted B. As such, reactions may occur as follows:



in which  $k_{ij}$  is the rate of reactions between the functional groups of monomer species  $i$  and  $j$ . We consider here only the irreversible regime, where the rate of bond breakage is negligible in comparison to the rate of bond formation. The step-growth copolymerization of monomers **A** and **B** is a second-order, bimolecular mechanism with reaction rates described respectively as:

$$\frac{-d[\text{A}]}{dt} = 2k_{\text{AA}}[\text{A}][\text{A}] + (k_{\text{AB}} + k_{\text{BA}})[\text{A}][\text{B}], \quad (3.1)$$

$$\frac{-d[\text{B}]}{dt} = 2k_{\text{BB}}[\text{B}][\text{B}] + (k_{\text{AB}} + k_{\text{BA}})[\text{A}][\text{B}], \quad (3.2)$$

where  $[A]$  and  $[B]$  are the instantaneous concentrations of unreacted binding moieties of **A** and **B**.

If we define  $[N]$  to be the total instantaneous concentration of unreacted binding moieties, such that  $[N] \equiv [A] + [B]$ , and take  $k_{AB} = k_{BA}$  by symmetry then we can combine Eq. 3.1 and Eq. 3.2 to yield:

$$-\frac{d[N]}{dt} = 2k_{AA}[A][A] + 4k_{AB}[A][B] + 2k_{BB}[B][B]. \quad (3.3)$$

In the case where the reaction starts with equal concentrations of monomer **A** and **B**, and the non-bonded attractions ( $\varepsilon_{ij}$ ) and activation energies ( $E_a^{ij}$ ) between **A** and **A** and between **B** and **B** are symmetric, then binding moieties **A** and **B** will be consumed at the same rate throughout the reaction. Accordingly, no composition drift is expected to occur under these conditions, and  $[N] = 2[A] = 2[B]$  will hold, on average, at any time during the polymerization. We note here that in working out the ideal kinetics behavior for the model in this section, we must limit our analysis to this symmetric interaction case, since  $E_a^{AA} \neq E_a^{BB}$  or  $\varepsilon_{AA} \neq \varepsilon_{BB}$  would impact the rate constants of the associated reaction pathways, leading to conditions where  $[A]_t \neq [B]_t$ . The resulting kinetics would require a more complex analysis and limit our ability to make comparisons with the numerical simulation results.

The change in the concentration of unreacted binding moieties can then be expressed as:

$$-\frac{d[N]}{dt} = k'[N][N], \quad (3.4)$$

where the effective rate constant for bond formation events in the system is:

$$k' = \frac{1}{2}(k_{AA} + 2k_{AB} + k_{BB}). \quad (3.5)$$

The expression in Eq. 3.4 may be integrated to obtain:

$$\frac{[N_0]}{[N]} = [N_0]k't + 1, \quad (3.6)$$

where  $[N]$  and  $[N_0]$  are the unreacted monomer binding moiety concentrations at time  $t$  and time  $t_0$ , respectively.

In terms of the concentrations of unreacted bonding moieties, the number average de-

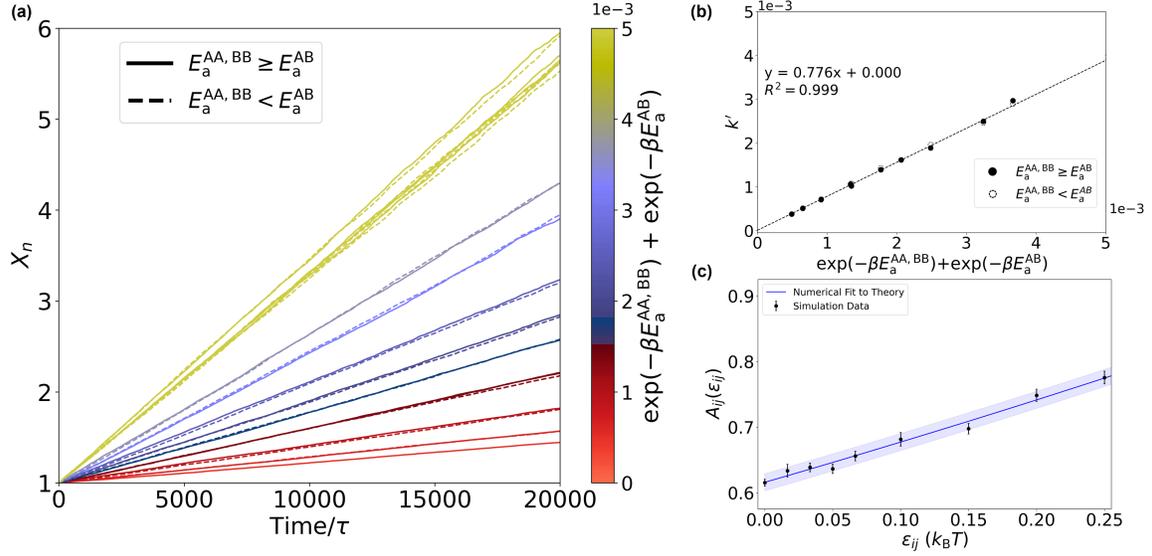


Figure 3.1: **Polymerization kinetics for low attraction strengths.** (a) Degree of polymerization,  $X_n$ , as a function of time at  $\varepsilon_{\text{all}} = 0.25k_B T$  for different combinations of activation energy between like-monomer reactions,  $E_a^{\text{AA, BB}}$ , and unlike-monomer reactions,  $E_a^{\text{AB}}$ . Coloration indicates the value of  $\{\exp(-\beta E_a^{\text{AA, BB}}) + \exp(-\beta E_a^{\text{AB}})\}$  for each case, and the solid and dashed line represent cases in which  $E_a^{\text{AA, BB}} \geq E_a^{\text{AB}}$  and  $E_a^{\text{AA, BB}} < E_a^{\text{AB}}$ , respectively. The exact set of  $E_a^{ij}$  combinations can be found in Table A.1 in Appendix A. (b) The effective rate constants  $k'$ , determined from the slope of each  $X_n$  curve in (a) via Eq. 3.8, are plotted as a function of  $\{\exp(-\beta E_a^{\text{AA, BB}}) + \exp(-\beta E_a^{\text{AB}})\}$ . Least-squares regression on the resulting values yields a linear relationship with a slope of 0.776 and a y-intercept of  $6e-7$ . The Arrhenius pre-exponential factor,  $A_{\text{all}}$ , may be determined from the slope of this regression line, according to Eq. 3.12. (c) The determined values of  $A_{ij}$  obtained from the process described in (a) and (b) are shown for several values of  $\varepsilon_{ij}$ . Each simulation data point corresponds to the result obtained from five simulation trials where  $\varepsilon_{\text{all}}$  was set equal to that value of  $\varepsilon_{ij}$ , with the error bars corresponding to the standard deviations between the trials. The resulting values were fit to the expected functional relationship of  $A_{ij}(\varepsilon_{ij})$  (blue line). The blue shaded region corresponds to the standard error in the predicted relationship of  $A_{ij}(\varepsilon_{ij})$ .

gree of polymerization,  $X_n$ , may be defined<sup>75</sup> as:

$$X_n \equiv \frac{[N_0]}{[N]}, \quad (3.7)$$

making Eq. 3.6:

$$X_n = [N_0]k't + 1. \quad (3.8)$$

As a result, we see that, for our simulated copolymerization,  $X_n$  is expected to grow linearly over time, as predicted by equal reactivity,<sup>73</sup> with the static rate constant of  $k'$  from Eq. 3.5.

**In simulations with low  $\epsilon_{\text{att}}$  values,  $X_n$  does increase linearly with reaction time, with a slope that depends on the activation energies.** Having obtained an expression for the growth of  $X_n$  within our model under ideal conditions, we next sought to compare our model to this expected behavior. To this end, we ran a series of simulations at low attraction strengths ( $\epsilon_{\text{all}} \leq 0.25k_B T$ ) – where in previous simulations we observed only ideal step-growth kinetics<sup>58,59,86</sup> – with different activation energy values for  $E_a^{\text{AA,BB}}$  and  $E_a^{\text{AB}}$ . [Note:  $E_a^{\text{AA,BB}} \equiv E_a^{\text{AA}} = E_a^{\text{BB}}$  at the symmetric limit where like monomers (**A:A** or **B:B**) interact equivalently, as discussed above.] The difference in the activation energies between like and unlike monomers spanned the range  $-4k_B T \leq \Delta E_a \leq 4k_B T$ , where  $\Delta E_a \equiv E_a^{\text{AA,BB}} - E_a^{\text{AB}}$ .

Figure 3.1a plots the development of  $X_n$  over time and shows that, for the full set of  $E_a^{ij}$  values explored,  $X_n$  increases linearly with reaction time, as expected from Eq. 3.8. These results show that, at low attraction strengths, our step-growth simulations behave ideally for a wide range of  $E_a^{\text{AA,BB}}$  and  $E_a^{\text{AB}}$  values, obeying second-order kinetics with a rate constant that is independent of the reaction progress.<sup>73</sup> Fig. 3.1a also shows that this rate behavior appears independent of the relative values of  $E_a^{\text{AA,BB}}$  and  $E_a^{\text{AB}}$ , as both cases with  $E_a^{\text{AA,BB}} \geq E_a^{\text{AB}}$  (solid lines) and cases with  $E_a^{\text{AA,BB}} < E_a^{\text{AB}}$  (dashed lines) show nearly identical rate behavior when the value of the quantity,  $\{\exp(-\beta E_a^{\text{AA,BB}}) + \exp(-\beta E_a^{\text{AB}})\}$ , indicated by the colorbar, is equal for both cases. In addition, the clear correspondence between the slopes of the lines in Fig. 3.1a and this quantity,  $\{\exp(-\beta E_a^{\text{AA,BB}}) + \exp(-\beta E_a^{\text{AB}})\}$ , suggests a connection between the effective rate constant in the system,  $k'$ , and this particular functional form of the activation energies between various monomer species. To explain this correspondence and the importance of this particular functional form, we will need to introduce the Arrhenius equation.

**At low  $\epsilon_{\text{att}}$  values, the dependence of the effective rate constant,  $k'$ , on the activation energies in the model can be fully explained using the Arrhenius formalism.** According to the Arrhenius equation,<sup>95</sup> the reaction rate constant depends on the activation energy,  $E_a$ , as follows:

$$k_{ij} = A_{ij} \exp(-\beta E_a^{ij}) \quad (3.9)$$

where  $\beta$  is the inverse temperature,  $1/(k_B T)$ , where  $T$  is the temperature and  $k_B$  is the Boltzmann constant, and  $A_{ij}$  is the Arrhenius pre-exponential factor between monomer

species  $i$  and  $j$ . Within collision theory  $A_{ij}$  depends upon the collisional cross-section, the reduced mass, and the steric factor for the specific pair of reacting species. Although the polymerization takes place in the solution phase, collision theory provides guidance on the factors that may influence  $A_{ij}$  in our simulations, especially given our lack of explicit solvent. In our model, both monomer species **A** and **B** have the same mass and are present in equal number. Further, they possess the same geometric structure, meaning there is no difference between them in terms of the orientational alignment needed for polymer bond formation. The only difference between **A** and **B** that can influence the value of  $A_{ij}$  lies in the strength of the pairwise attractive interactions between them,  $\epsilon_{ij}$ .

Plugging the Arrhenius equation into Eq. 3.5, we obtain an expression for the effective rate constant,  $k'$ , in terms of the activation energies,  $E_a^{ij}$ , for each reaction pathway:

$$k' = \frac{1}{2}[A_{AA}\exp(-\beta E_a^{AA}) + 2A_{AB}\exp(-\beta E_a^{AB}) + A_{BB}\exp(-\beta E_a^{BB})]. \quad (3.10)$$

When interactions and relative activation energies between like monomers are symmetric, as discussed above, then  $\epsilon_{AA} = \epsilon_{BB}$ . Given the equivalent **A** and **B** geometries and masses discussed above, this also means that  $A_{AA} = A_{BB}$ , which we term  $A_{AA,BB}$ . Therefore, at this symmetric limit, Eq. 3.10 simplifies to:

$$k' = A_{AA,BB}\exp(-\beta E_a^{AA,BB}) + A_{AB}\exp(-\beta E_a^{AB}). \quad (3.11)$$

From this point, we would like to determine how well the above equation holds in our simulations, while also determining the dependence of the Arrhenius pre-factor,  $A_{ij}$ , on the non-bonded attractions,  $\epsilon_{ij}$ , within our model. The simplest way to do so empirically is to run the polymerizations with the same non-bonded attractions acting between all monomer pairs, so that  $\epsilon_{AA,BB} = \epsilon_{AB}$ , which we will refer to as  $\epsilon_{\text{all}}$ . Since the prefactor  $A_{ij}$  only depends upon the monomer identities of  $i$  and  $j$  through its dependence on  $\epsilon_{ij}$ , then  $A_{AA,BB}(\epsilon_{\text{all}}) = A_{AB}(\epsilon_{\text{all}})$ , which we will refer to as  $A_{\text{all}}(\epsilon_{\text{all}})$ . Eq. 3.11 then becomes:

$$k' = A_{\text{all}}(\epsilon_{\text{all}})[\exp(-\beta E_a^{AA,BB}) + \exp(-\beta E_a^{AB})]. \quad (3.12)$$

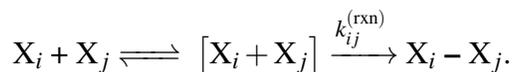
Although the polymerizations in Figure 3.1a were run at a range of  $E_a^{ij}$  values,  $\epsilon_{\text{all}} = 0.25k_B T$  for all cases, fixing the value of  $A_{\text{all}}$ . As a result, we now see from Eq. 3.12 that the slopes of the lines in Figure 3.1a ( $[N_0]k'$ , see Eq. 3.8) should increase linearly with the previously described quantity,  $\{\exp(-\beta E_a^{AA,BB}) + \exp(-\beta E_a^{AB})\}$ , as indicated by the

colorbar, given that  $[N_0]$  is held constant across the various simulation runs.

To test this prediction, in Figure 3.1b, we extract  $k'$  values from the results in Figure 3.1a by calculating the slopes of the  $X_n$  versus time lines there and dividing out the constant  $[N_0]$ . We then plot  $k'$  directly vs. the quantity  $\{\exp(-\beta E_a^{AA, BB}) + \exp(-\beta E_a^{AB})\}$  to obtain  $A_{ij}(\epsilon_{ij} = 0.25k_B T)$ . The results demonstrate a clear linear relationship, with an intercept of zero, in good agreement with the prediction of Eq. 3.12, indicating the validity of the Arrhenius formalism in describing the kinetic behavior of our model under ideal conditions.

**Determining the dependence of  $A_{ij}$  on  $\epsilon_{ij}$ .** According to Eq. 3.12, the slope of the line in Figure 3.1b provides the value of  $A_{\text{all}}$  at  $\epsilon_{\text{all}} = 0.25k_B T$ . Moreover, now that this linear relationship has been established, validating the applicability of Eq. 3.12 in our system, we calculate  $A_{ij}(\epsilon_{ij})$  in the same way for a series of  $\epsilon_{ij}$  values less than  $0.25k_B T$ , so that the non-bonded interactions between the monomers in these simulations are too weak to induce aggregation at the simulated oligomer lengths. Figure 3.1c shows our simulation results for these Arrhenius pre-exponential factors and how they change with  $\epsilon_{ij}$  under ideal second-order reaction conditions. Although the points in Figure 3.1c appear to follow a straight line, we do not expect the dependence of  $A_{ij}$  on  $\epsilon_{ij}$  to be linear. Since in the next section, we plan to use values of  $A_{ij}$  where  $\epsilon_{ij} > 0.25k_B T$ , we must be careful in extrapolating from the points we can calculate here under ideal kinetics conditions to those larger  $\epsilon_{ij}$  values where emergent oligomer assemblies alter those ideal kinetics. Determining the expected functional form involves considering the effect of  $\epsilon_{ij}$  on the relative number of closely interacting pairs of monomers within the system through a numerical integration of the radial distribution function and its dependence on  $\epsilon_{ij}$  through the inter-particle potential. Here we present an approach to identify the expected functional form of  $A_{ij}(\epsilon_{ij})$  based on the use of the radial distribution function  $g(r)$ , as adapted from the method outlined in Ref. [106].

We first consider the polymer bond formation process between monomers  $X_i$  and  $X_j$  under ideal conditions with a total effective rate constant of  $k_{ij}^{(\text{poly})}$ . We describe the formation of the polymer bond in two steps: (1) the formation of an interaction pair  $[X_i + X_j]$ , which is at equilibrium with its unpaired monomers, and (2) a subsequent irreversible reaction from a non-bonded interacting monomer pair to the bonded product with a rate constant of  $k_{ij}^{(\text{rxn})}$ :



The formation of the bonded end product, and thus the consumption of the monomer

species, depends entirely on the second step, enabling us to write the associated rate equation for the polymer bond formation as:

$$\frac{d[X_i - X_j]}{dt} = k_{ij}^{(\text{rxn})} [X_i + X_j] = k_{ij}^{(\text{poly})} [X_i] [X_j], \quad (3.13)$$

where  $\frac{d[X_i - X_j]}{dt}$  is the rate of formation of  $X_i - X_j$ , the polymerized product.

From Eq. 3.13, we can obtain an alternative expression for the effective rate constant  $k_{ij}^{(\text{poly})}$  for the pathway as:

$$k_{ij}^{(\text{poly})} = \frac{k_{ij}^{(\text{rxn})} [X_i + X_j]}{[X_i] [X_j]}. \quad (3.14)$$

The effective concentration of interacting pairs,  $[X_i + X_j]$ , can then be written<sup>106</sup> in terms of the radial distribution function between species  $i$  and  $j$ ,  $g^{(ij)}(r; \epsilon_{ij})$ , as:

$$[X_i + X_j] \approx 2\pi\rho_i\rho_j \int_0^{r'} g^{(ij)}(r; \epsilon_{ij}) r^2 dr, \quad (3.15)$$

where  $\rho_i$  and  $\rho_j$  are the number densities of species  $i$  and  $j$  respectively, and  $r'$  is the distance at which interactions between  $i$  and  $j$  become negligible. In our system, taking  $r' = r_{\text{cut}} = 2.5\sigma$  covers the full range of the modified Lennard-Jones potential (see Eq. 2.1 in Ch. 2). Eq. 3.15 therefore represents an approximation which captures the impact on the density of interacting pairs arising from the spatial correlations induced by the Lennard-Jones interaction. Combining Eq. 3.14&3.15, we obtain:

$$k_{ij}^{(\text{poly})} = 2\pi k_{ij}^{(\text{rxn})} \int_0^{r'} g^{(ij)}(r; \epsilon_{ij}) r^2 dr, \quad (3.16)$$

which is an expression for the effective rate constant for the pathway in terms of the integrated radial distribution function, which depends on  $\epsilon_{ij}$ , and the rate constant  $k^{(\text{rxn})}$ . As the rate determining step in the overall process, we expect  $k^{(\text{rxn})}$  to follow an Arrhenius relationship, that is:

$$k_{ij}^{(\text{rxn})} = A_{k^{(\text{rxn})}} e^{-\beta E_a^{ij}}, \quad (3.17)$$

where  $A_{k^{(\text{rxn})}}$  is the Arrhenius pre-exponential factor associated with this step. It is important

to note here that this definition considers the bond-formation rate from a collision complex,  $[X_i + X_j]$ , at an effective separation of  $r_0 = \sqrt[6]{2}\sigma$ , the minimum of the modified Lennard-Jones potential. At this distance, the activation energy depends only on the fixed parameter,  $\epsilon_{\text{rep}}$ , and not on the strength of the attractive portion,  $\epsilon_{ij}$ . We have therefore defined  $E_a^{ij}$  as the total potential energy of the reacting species above the minima at this distance, i.e., from the bottom of the LJ potential well. Given these definitions,  $A_{k(\text{rxn})}$  is independent of  $\epsilon_{ij}$ .

Further, as in Eq. 3.9, we assume that the overall rate constant  $k_{ij}^{(\text{poly})}$  is also Arrhenius, such that:

$$k_{ij}^{(\text{poly})} = A_{ij}(\epsilon_{ij})e^{-\beta E_a^{ij}}, \quad (3.18)$$

which contains our sought-after term  $A_{ij}(\epsilon_{ij})$ . Substituting the expressions of Eq. 3.17 into Eq. 3.16, and combining the result with Eq. 3.18 we obtain:

$$A_{ij}(\epsilon_{ij}) = 2\pi A_{k_{\text{bond}}} \int_0^{r'} g^{(ij)}(r; \epsilon_{ij}) r^2 dr. \quad (3.19)$$

The expression in Eq. 3.19 gives the desired value of  $A_{ij}$  in terms of constants independent of  $\epsilon_{ij}$ , and an integral of the radial distribution function, which contains the dependence on  $\epsilon_{ij}$  for  $A_{ij}$ .

The radial distribution function is a conditional probability distribution which describes the average density at a distance  $r$  from a given particle, relative to the bulk density.[107] It is related[107] to the potential of mean force  $w(r)$  as:

$$g(r) = e^{-\beta w(r)}, \quad (3.20)$$

where  $\beta = (k_B T)^{-1}$  is the inverse temperature for temperature  $T$  and Boltzmann constant  $k_B$ . The potential of mean force can be described as a combination of the pair potential governing the interactions between particles 1 and 1' (Eq. 2.1 in Ch. 2), and the reversible work done on the surrounding species when particles 1 and 1' are brought from  $r(1, 1') = \infty$  to  $r(1, 1') = r$ . In the dilute limit the reversible work is identically zero, and therefore for our system, we set:

$$g^{(ij)}(r) \approx e^{-\beta u_{\text{LJ}}(r(1, 1'), \epsilon_{ij})}, \quad (3.21)$$

where  $g^{(ij)}(r)$  is the radial distribution function for particle 1 of species  $i$  and particle 1' of species  $j$ .

Utilizing the radial distribution function, based on the expression in Eq. 3.21, and com-

binning with Eq. 3.19 we obtain:

$$A_{ij}(\epsilon_{ij}) = C_1 \int_0^{r_{\text{cut}}} e^{-\beta U_{\text{LJ}}(r; \epsilon_{ij})} r^2 dr = C_1 \sigma_{ij}(\epsilon_{ij}), \quad (3.22)$$

where  $r_{\text{cut}} = 2.5\sigma$  is the cutoff for the modified Lennard-Jones potential in our system,  $C_1$  is a collection of constants such that  $C_1 = 2\pi A_{k_{\text{bond}}}$ , and  $\sigma_{ij}(\epsilon_{ij})$  is a notational shorthand for the integral which we introduce for simplicity.

With Eq. 3.22 we have an expression which relates the Arrhenius pre-exponential factor  $A_{ij}$  to the strength of attractive Lennard-Jones interactions,  $\epsilon_{ij}$ , through the integral term  $\sigma_{ij}(\epsilon_{ij})$ . Our final step is to then determine the exact relationship between  $\sigma_{ij}(\epsilon_{ij})$  and  $A_{ij}(\epsilon_{ij})$  for our system.

Based on the arguments presented above, we expect that the functional dependence of  $A_{ij}(\epsilon_{ij})$  should be a simple linear transformation of the dependence of  $\sigma_{ij}(\epsilon_{ij})$ . In order to determine the correct dependence for our system, and thereby our sought after relationship for  $A_{ij}(\epsilon_{ij})$ , we compare values of  $\sigma_{ij}(\epsilon_{ij})$  calculated from Eq. 3.22 and values of  $A_{ij}(\epsilon_{ij})$  determined from simulation (see Eq. 3.12 and Fig. 3.1). We therefore perform simulations at small and fully equivalent values of  $\epsilon_{ij} = \epsilon_{\text{all}} \ll k_{\text{B}}T$ , where ideal kinetics hold and we may directly measure  $A_{ij}$  from the effective rate constant  $k'$  and the activation energies  $E_{\text{a}}^{ij}$ . For these same values of  $\epsilon_{ij}$ , values of  $\sigma_{ij}(\epsilon_{ij})$  were obtained via numerical integration of this term in Eq. 3.22, utilizing an adaptive Gaussian quadrature method.<sup>108</sup> Absolute error estimate for the obtained values were  $\sim 10^{-9}$  for all values of  $\epsilon_{ij}$  explored. A linear regression was then performed between the values of  $A_{ij}$  obtained from simulation and the values  $\sigma_{ij}$  obtained from integration at the same  $\epsilon_{ij}$ .

Fig. 3.2 shows the results of this analysis. The regression between  $\sigma_{ij}$  from Eq. 3.22 and  $A_{ij}$  from simulation are shown in the inset of Fig. 3.2 and clearly demonstrate the expected linear relationship between the two. The transformation between  $\sigma_{ij}(\epsilon_{ij})$  and  $A_{ij}(\epsilon_{ij})$ , as characterized by the regression equation in the inset, may be used for cases with higher  $\epsilon_{ij}$  in which ideal kinetic assumptions break down and  $A_{ij}$  is no longer directly measurable from the effective rate constant  $k'$ . In these cases, we can now determine the expected value for  $A_{ij}$  – had ideal kinetics still held – by calculating  $\sigma_{ij}(\epsilon_{ij})$  via numerical integration and applying the linear transformation obtained. The main plot of Fig. 3.2 shows the theoretical values of  $A_{ij}(\epsilon_{ij})$  obtained in this fashion, alongside those values obtained from simulation results. The theoretical values of  $A_{ij}$  shown here are utilized in Fig. 3.3 below for analyzing cases with attraction values for which ideal kinetics do not hold and  $A_{ij}$  cannot therefore be measured directly from simulation.

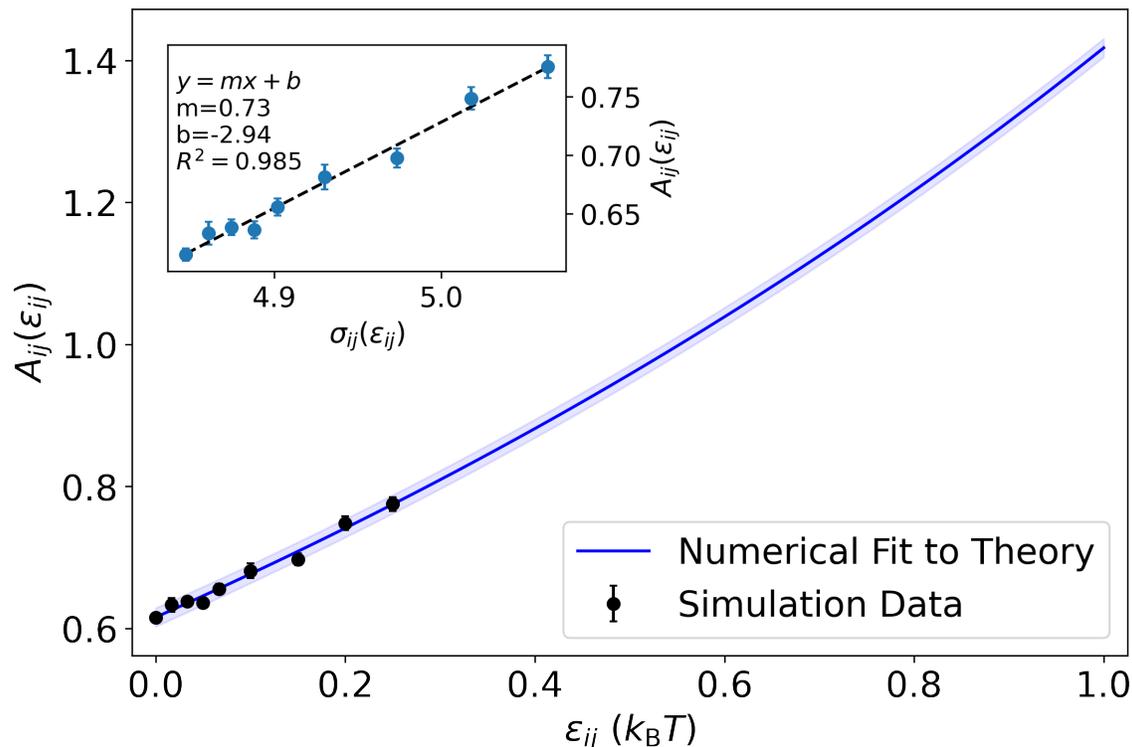


Figure 3.2: **Arrhenius pre-exponential factor as a function of  $\epsilon_{ij}$ .** The relationship between the pre-exponential Arrhenius factor  $A_{ij}$  and the attractive interaction strength  $\epsilon_{ij}$  is shown. Simulation data (black points) shows the mean value of  $A_{ij}(\epsilon_{ij})$  obtained from the observed effective rate constant  $k'$  for five separate simulation trials at chosen  $\epsilon_{ij}$ , with the error bars indicating standard deviation between the trials. The theoretical value (blue line) was obtained by calculating  $\sigma_{ij}$  from Eq. 3.22 and mapping the resulting  $\sigma_{ij}(\epsilon_{ij})$  to  $A_{ij}(\epsilon_{ij})$  using a linear regression. The shaded region indicates the standard error in the prediction obtained from the regression (see inset). **Inset:** A linear regression relating  $\sigma_{ij}(\epsilon_{ij})$  and  $A_{ij}(\epsilon_{ij})$ . Each point shows the mean value of  $A_{ij}(\epsilon_{ij})$  obtained from the effective rate constant  $k'$  (see Eq. 3.12) for five separate simulation trials at  $\epsilon_{ij} \leq 0.25k_B T$ , with the error bars indicating standard deviation between the trials. For each of these values of  $\epsilon_{ij}$ , the term  $\sigma_{ij}$  was obtained from Eq. 3.22 through a numerical integration. The resulting regression equation and fit parameters are shown.

As a result, we have now fully worked out the dependence of the effective rate constants in our model on both the non-bonded attractions, governed by  $\epsilon_{ij}$ , and the activation energies, governed by  $E_a^{ij}$ , such that we can write

$$k_{ij} = A_{ij}(\epsilon_{ij}) \exp(-\beta E_a^{ij}), \quad (3.23)$$

and use this to calculate the actual values of  $k_{ij}$  under conditions where the kinetics remain well-described by static reaction constants – that is, for  $\varepsilon_{ij} \leq 0.25k_{\text{B}}T$ .

### 3.3 Local sequence statistics depend on non-bonded attractions, activation energies, and emergent assemblies

In the preceding section, we showed how variations in non-bonded attractions and activation energies impact the polymerization kinetics in our model at low attractions via standard kinetics equations,<sup>73,75</sup> demonstrating that these simulations successfully reproduce such ideal kinetic behaviors under the relevant conditions. In this section, we examine how the combined effects of increased non-bonded attractions and varying activation energies can impact the sequences of the resulting co-oligomers – both under conditions where those standard kinetics hold and under conditions where emergent assemblies of nascent oligomers alter the reaction kinetics.

Enabled by our findings in the preceding section, we first determine an expression for the probability,  $p_{\text{AA,BB}}$ , of a having a bound, adjacent pair of like monomers (-**A-A**- or -**B-B**-) within the copolymer sequences under ideal kinetics. We then use this expression to investigate how our results align with or deviate from these predictions under conditions in which non-bonded attractions foster emergent heterogeneities in the reactant concentrations, as observed in our prior work for  $\varepsilon_{ij}$  values of about  $1k_{\text{B}}T$ .<sup>58,59,86</sup>

**Determining sequence pair probabilities as a function of  $\varepsilon_{ij}$  and  $E_a^{ij}$ .** We first consider the probability of a given **A** monomer in a chain subsequently binding to another **A** monomer. This conditional bonding probability,  $P(\mathbf{A}|\mathbf{A})$ , may be defined in terms of the rate constants of the accessible reaction pathways, namely  $k_{\text{AA}}$  and  $k_{\text{AB}}$ ; and may therefore be predicted<sup>98,109,110</sup> using:

$$P(\mathbf{A}|\mathbf{A}) = \frac{k_{\text{AA}}[\text{A}][\text{A}]}{k_{\text{AA}}[\text{A}][\text{A}] + k_{\text{AB}}[\text{A}][\text{B}]} \quad (3.24)$$

Dividing both the numerator and denominator by  $k_{\text{AB}}[\text{A}][\text{B}]$  yields:

$$P(\mathbf{A}|\mathbf{A}) = \frac{\frac{k_{\text{AA}}}{k_{\text{AB}}} \frac{[\text{A}]}{[\text{B}]}}{\frac{k_{\text{AA}}}{k_{\text{AB}}} \frac{[\text{A}]}{[\text{B}]} + 1} \quad (3.25)$$

As previously discussed, when interaction strengths and activation energies are sym-

metric between monomer species, no significant composition drift is expected to occur, and we can assume  $[A]_t = [B]_t$  for all time  $t$ . It is important to note here that this condition also implies that the probability of having an **A** or **B** monomer in the chain is equal. As a result, when no composition drift occurs, then equation 3.25 simplifies to:

$$P(\mathbf{A}|\mathbf{A}) = \frac{\frac{k_{AA}}{k_{AB}}}{\frac{k_{AA}}{k_{AB}} + 1}. \quad (3.26)$$

Similarly, the conditional probability of forming a B-B bond can be written as:

$$P(\mathbf{B}|\mathbf{B}) = \frac{\frac{k_{BB}}{k_{BA}}}{\frac{k_{BB}}{k_{BA}} + 1}. \quad (3.27)$$

Here it is helpful to make use of the reactivity ratios of Mayo-Lewis theory,<sup>74</sup> defining  $r_A \equiv \frac{k_{AA}}{k_{AB}}$  and  $r_B \equiv \frac{k_{BB}}{k_{BA}}$ , which allow Eq. 3.26&3.27 to be rewritten as:

$$P(\mathbf{A}|\mathbf{A}) = \frac{r_A}{r_A + 1}; \quad (3.28)$$

$$P(\mathbf{B}|\mathbf{B}) = \frac{r_B}{r_B + 1}. \quad (3.29)$$

Using the previously discussed Arrhenius<sup>95</sup> formalism for the relevant reaction rate constants, the reactivity ratios,  $r_A$  and  $r_B$ , can be related to the non-bonded attractions and the activation energies via:

$$r_A = \frac{A_{AA}}{A_{AB}} e^{-\beta(E_a^{AA} - E_a^{AB})}; \quad (3.30)$$

$$r_B = \frac{A_{BB}}{A_{BA}} e^{-\beta(E_a^{BB} - E_a^{BA})}, \quad (3.31)$$

where  $A_{ij}$  depends on  $\epsilon_{ij}$  as discussed in the previous section.

Here we again note that, because we have restricted our focus to cases where activation energies and Arrhenius pre-exponential factors are symmetric between like-monomer pairs **AA** and **BB**, then  $A_{AA} = A_{BB} \equiv A_{AA, BB}$  and  $E_a^{AA} = E_a^{BB} \equiv E_a^{AA, BB}$ . Similarly, for unlike monomer pairs **AB** and **BA**, we have, by symmetry,  $A_{AB} = A_{BA}$  and  $E_a^{AB} = E_a^{BA}$ . As a

result, we can write:

$$r_A = r_B = \frac{A_{AA, BB}(\epsilon_{AA, BB})}{A_{AB}(\epsilon_{AB})} e^{-\beta(E_a^{AA, BB} - E_a^{AB})}. \quad (3.32)$$

We now consider a random monomer within an oligomer sequence. The lack of composition drift, as discussed previously, means there is an equal probability that this monomer is either **A** or **B**. Thus the probability  $p_{AA, BB}$  of having a pair of like monomers next to one another in the sequence can be written as:

$$p_{AA, BB} = \frac{1}{2}P(\mathbf{A}|\mathbf{A}) + \frac{1}{2}P(\mathbf{B}|\mathbf{B}), \quad (3.33)$$

where the factor of one-half accounts for the probability that the conditional prior of having an **A** or **B** in the chain occurs.

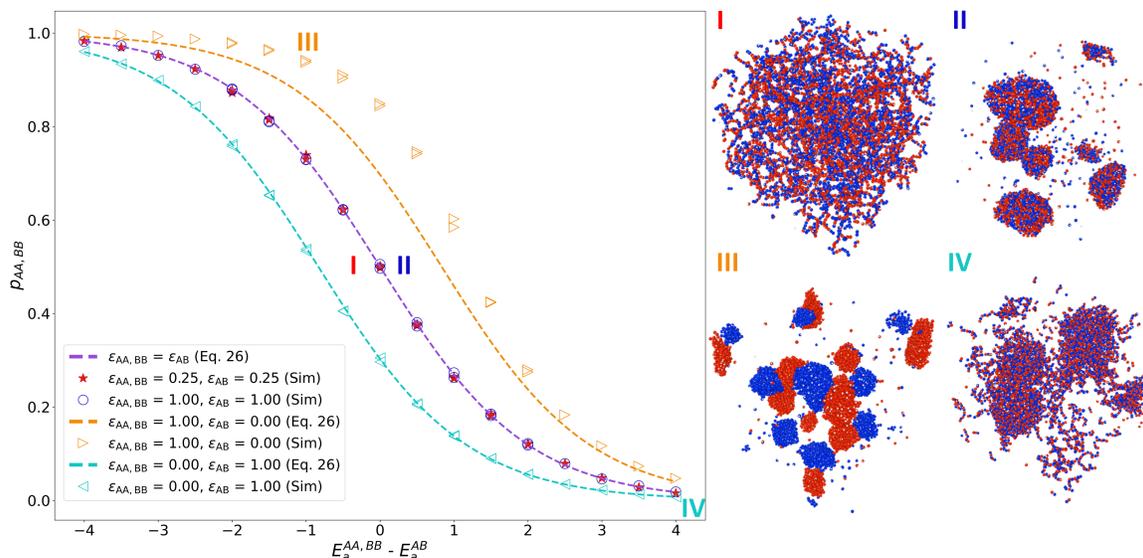
Finally, by combining the expression in Eq. 3.33 with the relations in Eq. 3.28- 3.32, we obtain:

$$p_{AA, BB} = \frac{\frac{A_{AA, BB}(\epsilon_{AA, BB})}{A_{AB}(\epsilon_{AB})} e^{-\beta(E_a^{AA, BB} - E_a^{AB})}}{\frac{A_{AA, BB}(\epsilon_{AA, BB})}{A_{AB}(\epsilon_{AB})} e^{-\beta(E_a^{AA, BB} - E_a^{AB})} + 1}. \quad (3.34)$$

This expression allows us to directly predict the probability of sequence pairs in the resulting copolymers generated at the ideal kinetics limit based on the difference in activation energies,  $E_a^{AA, BB} - E_a^{AB}$ , and the ratio of Arrhenius pre-exponential factors,  $A_{AA, BB}$  and  $A_{AB}$ , between like and unlike monomer species pairs in our system. These quantities are controlled through the model parameters  $E_{\text{barr}}^{ij}$  and  $\epsilon_{ij}$ , respectively, according to the dependencies outlined in Eq. 2.2 and Fig. 3.1c.

**The emergent assembly of nascent oligomers alters sequence when it alters the local concentration ratio of the comonomers.** Having obtained an expression for the probability of observing sequence pairs in the limit of ideal kinetic behavior, we now use this equation to investigate the combined effects of non-bonded attractions and activation energies on the sequence. We first validate the expression in Eq. 3.34 for systems which display the ideal kinetic behaviors under which it was derived. We then use it to explore the impact on the sequence of the emergent collective behaviors that cause these ideal kinetics assumptions to break down.

To this end, we explored four non-bonded attraction cases over a range of activation energies from  $-4k_B T \leq \Delta E_a \leq 4k_B T$ . The first,  $\epsilon_{\text{all}} = 0.25k_B T$ , is the same weak and symmetric attraction case for which we observed ideal kinetic behaviors in Fig. 3.1, and it



**Figure 3.3: System structures and sequence probabilities.** The probability of a matching sequence neighbor pair,  $p_{AA, BB}$ , is shown as a function of the activation energy difference  $E_a^{AA, BB} - E_a^{AB}$  for each of four different attraction strength combinations, as distinguished via marker style and coloration. Each point represents the value of  $p_{AA, BB}$  observed in simulation, averaged over three separate simulation trials for the associated simulation parameters. The dashed lines show the theoretical values of  $p_{AA, BB}$  which are obtained from Eq. 3.34, in which  $A_{ii}/A_{ij}$  is obtained from the relationship of  $A_{ij}(\epsilon_{ij})$  shown in Fig. 3.1c. The four structures shown on the right are representative system images taken at  $p = 0.9$ . These structures correspond to each of the four attraction strength combinations, at the points indicated by Roman numerals in the plot, with the colors corresponding to the data for each specific set of attractions.

thus serves as our validation case. We then explore the effect of higher attraction strengths for all monomers, setting  $\epsilon_{\text{all}} = k_B T$ . Finally, we examine cases of attractions strengths which differ between like and unlike monomers, testing both  $\epsilon_{AA, BB} = k_B T$ ,  $\epsilon_{AB} = 0$  and  $\epsilon_{AA, BB} = 0$ ,  $\epsilon_{AB} = k_B T$  combinations. For each attraction strength and  $E_a^{ij}$  combination, we simulated the copolymerization for three independent trials, collected the final sequences across these trials at reaction extent  $p = 0.9$ , and used them to calculate  $p_{AA, BB}$ , the probability of having a like neighbor (**A-A** or **B-B**) in the resulting sequence. Fig. 3.3 shows

the resulting values of  $p_{AA, BB}$  plotted as a function of  $\Delta E_a \equiv E_a^{AA, BB} - E_a^{AB}$  and compared these values (indicated with the points) to those predicted from the bonding probability calculated from Eq. 3.34 (indicated by the dashed lines).

For the two cases in which all monomers possess equal attraction strength, i.e.,  $\epsilon_{AA, BB} = \epsilon_{AB}$ , shown in red stars for  $\epsilon_{\text{all}} = 0.25k_B T$  and in blue circles for  $\epsilon_{\text{all}} = 1.0k_B T$ , there is nearly perfect agreement between these values and the predictions from Eq. 3.34 (dashed purple line). For these attractions, at  $E_a^{AA, BB} - E_a^{AB} = 0$ , there is an identical likelihood of finding like and unlike nearest neighbors in the resultant sequences, such that  $p_{AA, BB} = p_{AB, BA} = 0.5$ . The sequences which result under these conditions are fully random, as expected when there is no difference in either monomer interactions or activation energies across different monomer pairs. Decreasing  $E_a^{AA, BB} - E_a^{AB}$  increases  $p_{AA, BB}$ , and vice versa, and the changes in bonding probability with  $\Delta E_a$  agree with what is predicted via the Arrhenius formalism in Eq. 3.34. Notably, the sequence neighbor probability is independent of the value of  $\epsilon_{\text{all}}$ , as there is no discernible difference in  $p_{AA, BB}$  between the low attraction case (red stars) and the higher attraction case (blue circles) despite the fact that, at the higher attractions, when  $\epsilon_{AA, BB} = \epsilon_{AB} = k_B T$ , nascent oligomers aggregate into a more condensed phase – see Fig. 3.3, snapshot II.

For the case in which attractions act only between unlike monomers, namely  $\epsilon_{AA, BB} = 0$ ,  $\epsilon_{AB} = k_B T$ , we also observe sequence statistics (cyan, left-facing triangles) that are well predicted by Eq. 3.34 (cyan dashed line). Under these conditions, the bias in monomer attractions promotes the formation of unmatched pairs, lowering  $p_{AA, BB}$  as compared to the case where  $\epsilon_{AA, BB} = \epsilon_{AB}$ . This reduction is an expected consequence of Eq. 3.34, and the observed values of  $p_{AA, BB}$  are consistently well predicted across the full range of activation energy values explored.

In contrast, when attractions act only between like monomers in Fig. 3.3, the predicted sequence pair probabilities (orange dashed line) significantly underestimate the observed proportion of sequence pairs (orange, right-facing triangles). Under these conditions, as polymerization proceeds, like-monomer attractions drive an emergent phase separation into **A**-rich and **B**-rich regions (Fig. 3.3, snapshot III). This phase separation locally enriches like-monomer concentrations and produces a significant increase in the likelihood of like sequence pairs compared to the prediction from Eq. 3.34, which does not take into account emergent concentration heterogeneities and, as a result, underestimates  $p_{AA, BB}$  by  $> 0.14$  in some cases.

The distinct effects of attraction-biasing on phase behavior in each of these cases is crucial to their respective influence on sequence. For non-bonded attractions between only

unlike monomers, the system condenses into an aggregate phase as the polymerization proceeds, however, the comonomers within the aggregated phase remain well-mixed. Thus, though the local concentration of reactants is higher than the initial bulk concentrations from which the predictive kinetics were derived, it remains homogeneous with respect to the comonomer reactant species. In contrast, when attractions occur between only like monomers, the comonomers segregate into different regions of the aggregated phase, and it is the resulting heterogeneity that causes the breakdown in the predictive capability of Eq. 3.34.

It is important to note, however, that, as predicted, both the attractive interactions and activation energies are key to the final sequence statistics. In all four attraction combinations explored, a sufficient imbalance in the activation energies will drive the formation of predominantly "blocky" ( $E_a^{AA, BB} < E_a^{AB}$ ) or predominantly alternating ( $E_a^{AA, BB} > E_a^{AB}$ ) sequences, even when the attractive interactions alone would promote the opposite sequence statistics. The impact of the emergent assembly behaviors driven by non-bonded attractions on sequence is most apparent in cases where  $E_a^{AA, BB} \approx E_a^{AB}$ , when the biasing due to activation energy is minimal and sequences would otherwise be nearly random. The results in Fig. 3.3 demonstrate that non-bonded attractions and activation energies yield combined effects on sequence in a manner than cannot be fully captured by conditional bonding probabilities derived from ideal kinetic behaviors.

### 3.4 Differing attraction strengths and activation energies can drive long-range sequence biasing

In our discussion of sequence biasing to this point, we have considered only the statistics of nearest neighbors in sequence space through the observed probabilities of specific sequence pairs. We now consider the extent to which the biasing we observe extends throughout the sequence and how it is impacted by activation energies and attractive interactions.

To examine long-range correlations in the final oligomer sequences, we employ a previously developed metric<sup>111</sup> which maps the copolymer sequence to a 1D random walk, with steps of +1 and -1 for the **A** or **B** monomers, respectively. The root-mean-square fluctuation,  $F(l)$ , is defined as:

$$F(l) \equiv \sqrt{x(l)^2 - \overline{x(l)}^2}, \quad (3.35)$$

where  $x(l)$  is the total displacement along a sequence walk of  $l$  steps, and the overbars represent an averaging over all sequences in the system and all starting points within a sequence. In general,  $F(l) \propto l^\alpha$ , with  $\alpha = 0.5$  for a random walk. A fluctuation metric that scales with  $\alpha > 0.5$  indicates correlation between monomers in the sequence, corresponding to "blocky" sequences, while scaling of  $\alpha < 0.5$  indicates anti-correlation, corresponding to alternating sequences. Such correlations persist over a range  $R$ , such that, when  $l > R$ , the random scaling behavior ( $\alpha = 0.5$ ) is recovered. Thus, by comparing the scaling behavior of  $F(l)$  for a set of sequences from the simulated copolymerizations to the known scaling of  $F(l)$  for a fully random walk, it is possible to determine not only the presence of long-range correlations and anti-correlations in the sequence, but also the sequential distance over which these correlations persist.

Fig. 3.4 shows the results of this metric applied to each of the four  $\epsilon_{ij}$  combinations presented in Fig. 3.3 over the full range of activation energies, such that  $-4k_B T \leq \Delta E_a \leq 4k_B T$ . In each case, the scaling for a random sequence walk where  $F(l) \propto l^{1/2}$  is shown by the dashed black line. For attractive interactions where  $\epsilon_{AA, BB} = \epsilon_{AB}$  (Fig. 3.4, top row), the behavior of  $F(l)$  depends strictly upon  $\Delta E_a$ , the difference between the activation energies in the system, as the results for  $\epsilon_{\text{all}} = 0.25k_B T$  are indistinguishable from the results for  $\epsilon_{\text{all}} = k_B T$ . In both of these cases, the sequences at  $\Delta E_a = 0$  (red line) reproduce the expected random sequence scaling with  $\alpha = 0.5$ .

As this balance shifts and  $E_a^{\text{AA, BB}} < E_a^{\text{AB}}$ , the slope of  $F(l)$  increases noticeably, showcasing the bias towards blocky sequences induced by these combinations of activation energies. Importantly the magnitude of  $\Delta E_a$  impacts not only the extent of the biasing, but also the range over which it is observed. For small differences in the activation energies between species pairs, the increase in  $\alpha$  is comparably small, and persists only for the first few monomers along the chain before the  $\alpha = 0.5$  behavior returns. As this difference in activation energy decreases further and  $\Delta E_a = -4k_B T$ , sequence correlation is larger and persists beyond  $l = 10$ , the average degree of polymerization in the system, indicative of highly blocky sequences that are correlated throughout. Analogous behaviors are seen when the activation energies instead bias the system towards alternating sequences, i.e., when  $E_a^{\text{AA, BB}} > E_a^{\text{AB}}$ . Small positive  $\Delta E_a$  values yield an anti-correlation only for their nearest neighbors ( $l = 1$ ) before reverting to the random sequence scaling. Further increasing  $\Delta E_a$  results in sequences that are almost entirely alternating, and the corresponding oscillation in  $F(l)$  is readily apparent, persisting beyond  $l = 10$ .

The interplay between the effects of non-bonded attractions and activation energies in influencing sequence behavior may be seen in cases for which  $\epsilon_{AA, BB} \neq \epsilon_{AB}$  in Fig. 3.4,

bottom row. As in the cases with equivalent attractions between all monomers, sequence statistics ranging from predominantly alternating to predominantly blocky may be induced by the balance of activation energies in the system. Unlike the equivalent attraction cases, however, the activation energy combinations required to reach each of these regimes differs, as the bias arising from activation energy may act contrary to, or in concert with, the sequence biasing caused by the differences in attraction. When attractions act only between

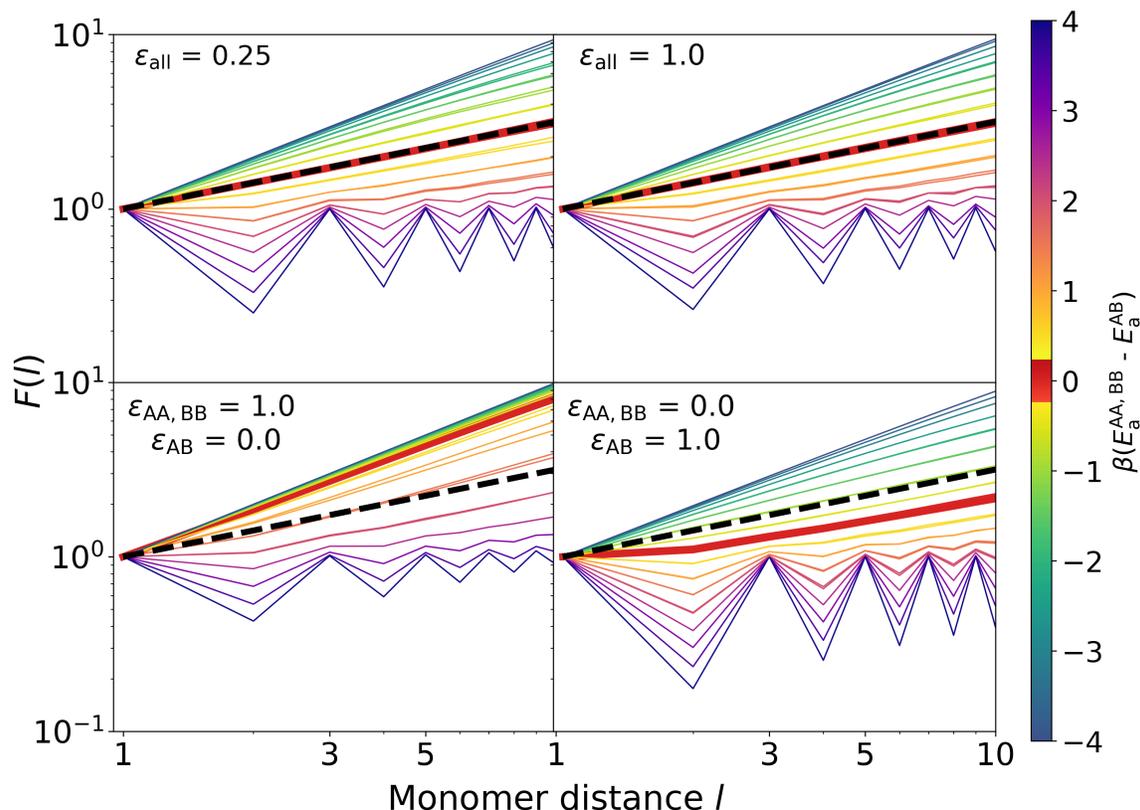


Figure 3.4: **Long-range sequence statistics.** The root-mean-square fluctuation in sequence,  $F(l)$ , as a function of monomer distance  $l$ , is shown for four different attractions:  $\epsilon_{all} = 0.25 k_B T$ ;  $\epsilon_{all} = k_B T$ ;  $\epsilon_{AA, BB} = k_B T$ ,  $\epsilon_{AB} = 0$ ; and  $\epsilon_{AA, BB} = 0$ ,  $\epsilon_{AB} = k_B T$ . Coloration indicate the values of  $\beta E_a^{AA, BB} - \beta E_a^{AB}$  for each curve. The sequence fluctuation metric is calculated via  $F^2(l) = \overline{x(l)^2} - \overline{x(l)}^2$ , where  $x(l)$  is the displacement in a sequence walk of  $l$  steps away from the initial sequence position.<sup>111</sup> The dashed black line indicates the fluctuation metrics expected for random sequences, where  $F(l) = l^\alpha$  and  $\alpha = 0.5$ . The sequences were obtained from three independent simulations trials for each parameter combination.

like monomers (Fig. 3.4, bottom left), the case where  $\Delta E_a = 0$  (red line) shows significant and long-range sequence correlation. In addition, for some cases where  $E_a^{AA, BB} > E_a^{AB}$  the effect of the non-bonded attractions overpowers the competing influence of  $\Delta E_a$ , resulting in correlated, blocky sequences. Only when the activation energy difference is sufficiently large,  $\Delta E_a \gtrsim 2k_B T$ , do the anti-correlations and  $F(l)$  oscillations indicative of predominantly alternating sequences emerge, and even then not to the extent seen for the  $\epsilon_{AA, BB} = \epsilon_{AB}$  cases. For attractions only between unlike monomers, (Fig. 3.4, bottom right), the opposing effect is observed, where the  $\Delta E_a = 0$  case (red line) shows short-range anti-correlation, with long-range correlation and blocky sequences only emerging for  $\Delta E_a^{AA, BB} \lesssim -2k_B T$ . Thus, both the relative attractions and activation energies contribute to the sequence development of oligomers, and imbalances in either or both of these quantities can lead to a broad range of sequence statistics spanning the spectrum between alternating and blocky copolymers.

### 3.5 Activation energies and non-bonded attractions influence block and chain length distributions

In the preceding sections we have demonstrated how differences in non-bonded attractions and activation energies between monomer species can lead to changes to the kinetic and sequence behaviors, producing features not predicted or observed under ideal kinetics conditions. Here we consider how the combined influence of the altered kinetics and sequence biasing impacts the structure of the resulting oligomers. To this end, we examine the distribution of chain lengths and block lengths within our system and compare these to expected distributions from Flory-Schulz and Markovian statistics under select combinations of monomer attractions and activation energies. Additionally, we consider how these distributions may be further impacted by the stiffness of the resulting oligomer chain.

The expected distribution of chain lengths for a step-growth polymerization under ideal conditions follows the well-known Flory-Schulz distribution,<sup>112</sup> given by:

$$P(x) = (1 - p)p^{x-1}, \quad (3.36)$$

where  $p$  is the reaction extent and  $x$  is the length of the chain. If the sequences within these chains develop according to a Markov chain model, in which only the influence of the preceding monomer impacts the probability of the next monomer in sequence, then the

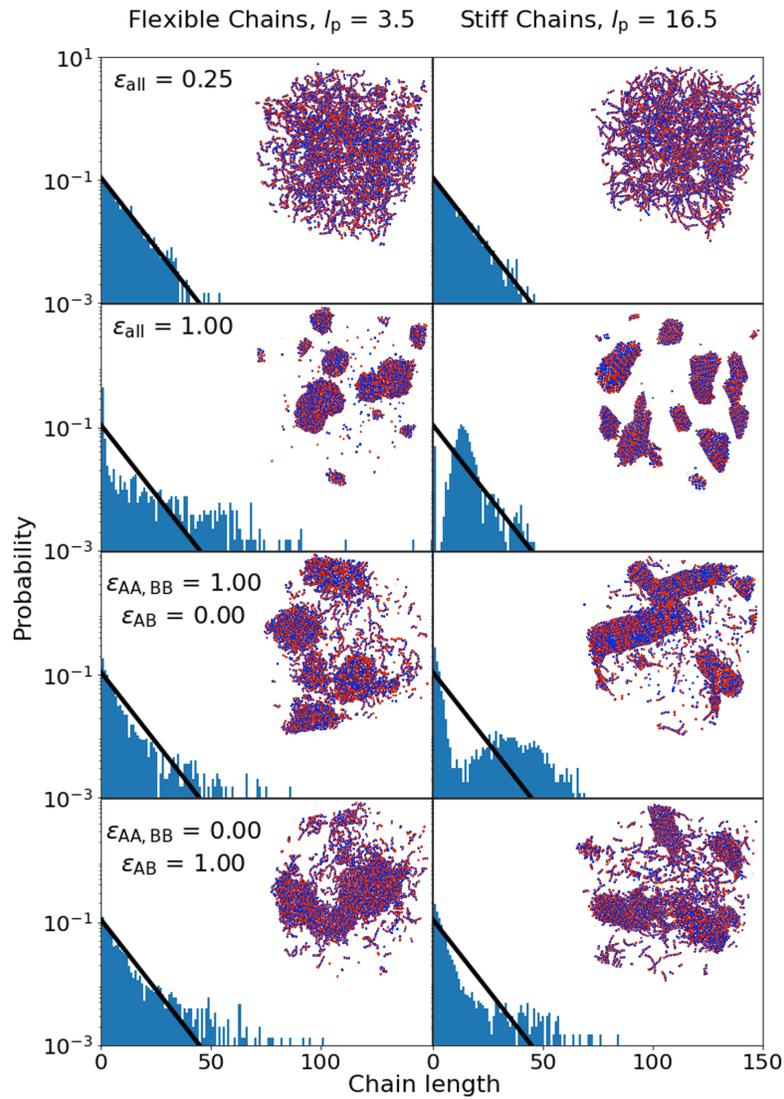


Figure 3.5: **Chain-length distributions and system structures.** The distribution of chain lengths are shown for four different attraction strengths at  $E_a^{AA, BB} - E_a^{AB} = 2 k_B T$ . Each row corresponds to a different attraction strength combination, as indicated by the label. The left column shows the results for flexible chains ( $l_p = 3.5$ ) and the right column for stiffer chains ( $l_p = 16.5$ ). The solid back lines show the predicted Flory-Schulz distribution at  $p = 0.9$ , as calculated from Eq. 3.36. For each combination of system parameters, a representative system structure taken at  $p = 0.9$  is inset alongside the distribution.

expected block length distribution can be written<sup>58</sup> analogously as:

$$P(n) = (1 - p_{AA, BB}) (p_{AA, BB})^{n-1}, \quad (3.37)$$

where  $p$  is the reaction extent,  $n$  the block length, and  $p_{AA, BB}$  the probability of an **AA** or **BB** sequence pair occurring. The expression in Eq. 3.36 depends on the condition of a fixed monomer reactivity throughout the reaction, and that of Eq. 3.37 on the condition that there are no long-range effects influencing sequence development. Both these conditions break down under certain activation energies and monomer attractions in our simulations, and the degree to which the ideal distributions in Eq. 3.36&3.37 are reproduced reports on how well these underlying assumptions hold for specific cases.

Fig. 3.5 shows chain length distributions observed at  $p = 0.9$  for each of the attractive interaction cases previously explored, when  $\Delta E_a = 2k_B T$ . In each of these cases, additional simulations were run with increased chain stiffness, corresponding in an increased persistence length from  $l_p = 3.5$  to  $l_p = 16.5$ , which are shown alongside the standard, flexible chain conditions. In all cases, the Flory-Schulz distribution (black line) and a representative system structure for each case at  $p = 0.9$  are also shown. The top row of Fig. 3.5 shows the case of  $\epsilon_{\text{all}} = 0.25k_B T$ , which reproduces ideal kinetic behaviors in Fig. 3.1. For both flexible and stiff chains, the chain length distribution observed is in good agreement with the predicted Flory-Schulz distribution, and no sign of oligomer aggregation is apparent in the resulting structure. As the attractions are increased to  $\epsilon_{\text{all}} = k_B T$  in the second row of Fig. 3.5, there is a clear formation of aggregates which coincides with a deviation from the Flory-Schulz distribution. Notably, the type of deviation observed is also influenced by the chain stiffness, with flexible chains showcasing a general broadening of the distribution towards greater chain lengths, while stiff chains demonstrate a sharply peaked distribution. This type of peaked distribution for stiff chains was observed in prior work<sup>59</sup> and is indicative of a characteristic chain length which emerges as the reaction proceeds, a consequence of nematic alignment in the system.

The peaked nature of the distribution for stiff chains is a consequence of increased chain stiffness leading to nematic alignment in neighboring oligomer chains,<sup>59</sup> increasing the collective effect of the non-bonding attractions within such aligned chains. The consequences can be seen in the trend in peak behavior for stiff chains down the right-hand column of Fig. 3.5. When attractions are strong enough for emergent aggregates to form, differences in interaction strengths between monomer species have important effects on both the structure of the aggregates and the associated shift in chain length distribution.

When  $\epsilon_{\text{all}} = k_{\text{B}}T$ , nascent chains can readily align irrespective of the sequence properties of neighboring chains, which produces more numerous and smaller aggregates, with almost complete monomer incorporation by the time 90% of functional groups have reacted, since monomers can readily incorporate anywhere, regardless of the nearby aggregate composition. This ease of monomer incorporation results in a distribution that has almost no monomers and dimers and is sharply peaked. In contrast, when attractions act only between like-monomers ( $\epsilon_{\text{AA, BB}} = k_{\text{B}}T, \epsilon_{\text{AB}} = 0$ ), monomers do not incorporate as easily into aggregate regions enriched in the other monomer species. As a result, the chain length distribution and structural snapshot show more free monomers, dimers, and other short chains at  $p = 90\%$ , coupled with a smaller number of larger aggregates with a wider chain length distribution. When attractions act only between unlike monomers ( $\epsilon_{\text{AA, BB}} = 0, \epsilon_{\text{AB}} = k_{\text{B}}T$ ), this effect is even stronger, with an even greater proportion of the distribution shifting towards short chains, less complete aggregation in the snapshot, and an even broader tail of longer chains.

It should be noted that these trends do not operate independently of activation energy. While Fig. 3.5 shows results for cases where  $E_{\text{a}}^{\text{AA, BB}} > E_{\text{a}}^{\text{AB}}$ , the relative abundance of matched pairs in the sequences will be increased when activation energies are instead balanced ( $\Delta E_{\text{a}} = 0$ ) or imbalanced so as to promote the formation of matched pairs ( $E_{\text{a}}^{\text{AA, BB}} < E_{\text{a}}^{\text{AB}}$ ). As the relative abundance of like sequential pairs increases, the capacity of collective monomer attractions to drive chain alignment increases and thereby strengthens the peaked nature of the chain length distribution (see Fig. A.5 in Appendix A).

The results in Fig. 3.5 confirm that the predictions of the Flory-Schulz distribution in Eq. 3.36 only hold for cases that remain homogeneous throughout the reaction. They demonstrate that emergent heterogeneity, caused by polymerization induced aggregation due to non-bonded attractions, causes a breakdown in the founding assumptions in Flory-Schulz and fundamentally alters the distribution of chain lengths. For all cases in which aggregates form upon polymerization, we observe non-standard kinetic behaviors and shifts in the chain length distributions away from the ideal Flory-Schulz behavior. The precise nature of the shift depends upon the nature of this aggregated phase and how it is impacted by both the stiffness of the nascent oligomer chain, which impacts the morphology of the aggregate structures, and the relative non-bonded attractions and activation energies between the comonomer species, which determines the relative ability of monomers and small oligomers to incorporate into the aggregate phase.

For the analogous expression for the distribution of block lengths in Eq. 3.37, a similar shift away from the Markovian distribution occurs when the emergent assembly produces

the type of long-range sequence correlation seen in Fig. 3.4. These correlations are strongest when attractions are only between like monomers ( $\epsilon_{AA, BB} = k_B T, \epsilon_{AB} = 0$ ), which promotes the formation of blocky sequences. In Fig. 3.6, we show the block length distributions for contiguous all-**A** or all-**B** blocks at select conditions with these attractions. Three different activation energy combinations are shown, which, from top to bottom, bias the sequences towards alternating, random, and blocky sequences, respectively. Results are shown for both flexible and stiff chain cases. In each case, the block length distribution (purple) is shown atop the chain length distribution (grey). The Markovian predicted distribution of block lengths from Eq. 3.37 (black line) is also plotted in each case. This distribution is calculated using values for  $p_{AA, BB}$  that are predicted based on the ideal kinetic behavior (Fig. 3.3 dashed lines), which allows us to characterize the shift in block distribution due to both the nearest neighbor (Fig. 3.3) and long-range (Fig. 3.4) sequence biasing arising from a breakdown in ideal kinetic behaviors.

When  $\Delta E_a = 2k_B T$  (Fig. 3.6, top row) the bias in activation energy towards unlike monomer pairs is sufficient to overcome the biasing of attractions towards like monomer pairs, and only very short sequence blocks are observed. For this case, the only sign of sequence biasing was an anti-correlation in nearest sequence neighbors (Fig. 3.4 lower-left panel). The underlying presumption of Markov statistics, that only nearest sequence neighbors impact the probability of the next monomer in the sequence, therefore holds. Indeed, the expression in Eq. 3.37 derived from this assumption shows good agreement with our observed block length distribution in this case.

For both  $\Delta E_a = 0$  (Fig. 3.6, middle row) and  $\Delta E_a = -2k_B T$  (Fig. 3.6, bottom row), we observed long-range correlations in sequence statistics in Fig. 3.4 (lower-left panel), and here the underlying assumption of a Markov chain and the associated prediction from Eq. 3.37 clearly break down. In these cases, the activation energy biasing is either neutral ( $\Delta E_a = 0$ ) or works in concert with attraction biasing ( $\Delta E_a = -2k_B T$ ), producing blocky sequences with contiguous segments well beyond what is predicted from the Markovian assumption.

For stiff chains, the same type of sharply peaked behavior seen in the chain length distribution is reflected in the block length distribution, suggesting that the characteristic length scale of stiff chain aggregates is impacting block length. For flexible chains no such peaks are observed, and the distribution is generally broadened towards greater block lengths, matching the behavior of the chain length distribution. Notably, for flexible chains with  $\Delta E_a = -2k_B T$ , the combined biasing of activation energy and like monomer attractions becomes so extreme that the block distribution shows a near complete correspondence to

the chain length distribution – that is the resulting oligomers are largely homopolymers of **A** or **B**, with **A–B** contacts primarily occurring at chain ends or at the block interface of the longest chains, similar to structure III in Fig. 3.3. These conditions may be regarded as nearing an “incompatibility” limit for the copolymerization of **A** and **B**, beyond which the system behaves analogously to separate, but simultaneous, homopolymerizations of **A** and **B**.

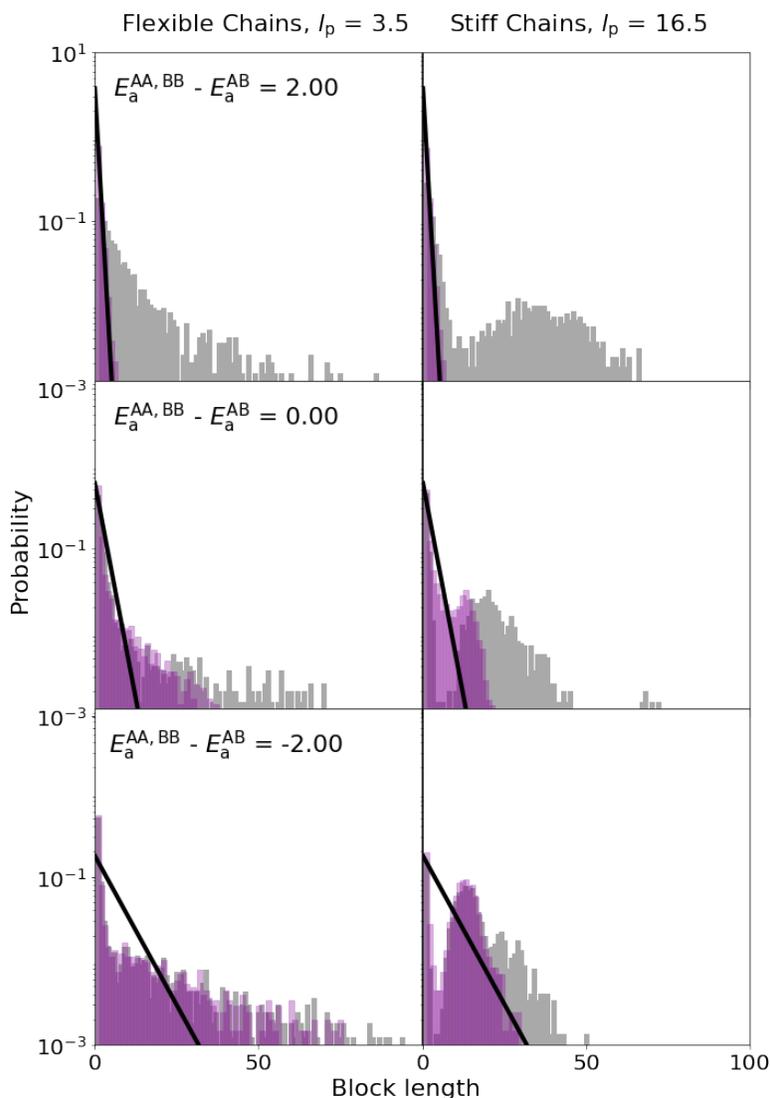


Figure 3.6: **Block-length distributions.** The distributions of block lengths are shown here for systems where  $\epsilon_{AA, BB} = k_B T$  and  $\epsilon_{AB} = 0$ . Each row corresponds to a different combination of activation energies,  $\Delta E_a$ , as indicated by the label. The left column shows results for flexible chains ( $l_p = 3.5$ ) and the right column for stiffer chains ( $l_p = 16.5$ ). For each combination of system parameters shown, the block length distribution (purple) is shown against the chain length distribution (grey). The solid black lines indicate the expected Markovian statistics (from Eq. 3.37) calculated from the predicted sequence neighbor probability  $p_{AA, BB}$  (from Eq. 3.34 and Fig. 3.3).

### 3.6 Summary and conclusions

In this Chapter, we have examined the change in kinetic behaviors in a solution based A,B-copolymerization in response to changes in inter-monomer attractions and in the activation energies of the different reaction pathways. Comparing the results to standard kinetic theories of step-growth polymerization, we find that such theories are effective in describing the kinetics of reactions which remain homogeneous throughout. However, in keeping with previous studies,<sup>58</sup> we find that standard kinetics break down when non-bonded attractions between reactants lead to an emergent concentration heterogeneity in the solution. The polymerization-driven assembly of nascent oligomers alters the kinetics and sequence development of the reaction by altering the local concentration ratio of the comonomers. When this emergent concentration heterogeneity happens in polymerizations with differing comonomer attractions and reactivities, a diverse set of sequence statistics results, ranging from highly block-like to almost exclusively alternating sequences. Depending on the non-bonded attractions between different comonomer pairs in the system, the bias introduced by assembly may act to reduce sequence biasing arising from reactivity differences, yielding more random sequences, or to further reinforce reactivity biasing, yielding more ordered sequences. Thus, the sequence statistics which result when emergent phase separation occurs depends on a complex interplay between non-bonding interactions, comonomer reactivity, and even the geometric properties of the resulting oligomers.

Limiting our kinetic analysis to symmetric interaction cases that proceed without composition drift enables us to demonstrate the magnitude of the deviations from standard kinetic behaviors that can result from the emergent phase separation we observe. Under asymmetric interaction conditions in which compositional drift does occur, similar assembly phenomenon are also expected to occur and influence reaction kinetics and sequence,<sup>86</sup> which will be explored in Chapter 4. The degree of this deviation is also sensitive to both the extent of the reaction and the magnitude of the attractions driving assembly, with attractions  $> k_B T$  leading to an earlier emergence of concentration heterogeneity, and even very weak  $< k_B T$  attractions capable of producing phase separation later in the reaction once oligomers reach sufficient length.<sup>58,86</sup> Thus, the emergence of collective behaviors among the comonomers and nascent oligomers is expected to influence the resulting sequences of statistical copolymers for a wide variety of effective comonomer attractions and reactivities. These results showcase the potential to exploit polymerization induced assembly behaviors for sequence control in the design of advanced copolymeric materials. They also suggest that the informed management of the relative non-bonded interactions and

activation energies between different comonomers, through comonomer selection, chemical modification, or solvent choice, is a potential avenue for efficiently biasing sequences within one-pot syntheses.

## Chapter 4

# Reactant Assembly In Selective Solvent Conditions

### 4.1 Introduction

In prior work,<sup>58,59</sup> and in the preceding Chapter,<sup>87</sup> we investigated only monomer attractions that act symmetrically between “like” monomers (i.e., between all monomers of the same type) and showed that these can give rise to polymerization-induced aggregation and demixing between the monomer types, which then biases additional bond formation and shifts the overall copolymer sequence. However, many monomer combinations would be expected to have asymmetric interactions, where one type of monomer has a stronger or weaker effective self-attraction than the other type, due either to the nature of the monomer-monomer interactions themselves, or to the influence of a solvent selective for one of the monomer species.<sup>113–115</sup> Indeed, such an asymmetry in the effective monomer-monomer interactions, as mediated through a selective solvent, is a crucial feature of PISA formulations. The solvophilic stabilizing blocks in these chain growth polymerizations allow for the addition of a solvophobic monomer at concentrations that would otherwise not be well solvated.<sup>66,67</sup>

In this Chapter, we consider the case of a selective solvent, utilizing an extension of

---

This Chapter has been adapted with minor alterations from Ref. [86] with permission from the Royal Society of Chemistry. Ryan L. Hamblin contributed to the investigation, software, formal analysis, visualization, and writing. Nhu Q. Nguyen contributed to the data curation, visualization, and editing. Kateri H. DuBay contributed to the conceptualization, methodology, supervision, formal analysis, and writing.

our coarse-grained, reactive model for step-growth polymerization. The copolymerization between solvophobic and solvophilic species is modeled by assigning stronger effective attractions for the more solvophobic monomer and weaker attractions for the more solvophilic monomer, as discussed in Chapter 2. In Section 4.2 we examine the kinetic behaviors of copolymerization under such asymmetric attractions. In Section 4.3 we investigate the sequences of the resulting oligomers, and the influence of attraction asymmetry on sequence. We analyze the impact of chain stiffness in Section 4.4 and discuss the aggregation behavior of the resultant oligomers in Section 4.5. We find that selective solvation during polymerization can drive an emergent phase separation, yielding a complex interplay between the sequence and supramolecular assembly of the nascent oligomers.

## 4.2 Solvent mediated attractions can drive an emergent assembly process, altering reaction kinetics

We first aim to explore whether collective monomer and nascent oligomer behaviors would lead to changes in the local solution environment in cases with asymmetric solvent affinities, analogous to the interactions that give rise to bootstrap effects and PISA syntheses. To this end we initially sought to understand how well the self-association of the more strongly self attractive **A** monomers and less self attractive **B** monomers drives the formation of local regions of concentration enrichment and how those associations influence reaction kinetics and the resulting system dispersity. We observe that for effective attractive interactions of sufficient strength, an emergent self organization of monomers and oligomers occurs, which alters the solution structure and produces unconventional kinetic features. Fig. 4.1a shows the number-averaged degree of polymerization,  $X_n$ , as a function of reaction time for the full range of attractive interaction strengths studied. There are two important observations regarding the reaction kinetics displayed. First, at lower attraction strengths,  $X_n$  increases linearly at a rate that is independent of the reaction progress. This linearity is expected for a second-order reaction scheme and has been experimentally observed in step-growth polymerizations.<sup>73</sup> Thus, for these lower attraction strengths, our model successfully captures the kinetics of a step-growth polymerization based on Flory's equal reactivity principle,<sup>73</sup> and we observe a fixed reaction rate irrespective of oligomer length. Second, we note that the reaction rate actually increases over time for cases with stronger solvent-mediated interactions, i.e., higher values of  $\Sigma \epsilon_{ij} \equiv \epsilon_{AA} + \epsilon_{AB} + \epsilon_{BB}$ . In these cases, a transition occurs at some point when the rate of the reaction increases, after which  $X_n$

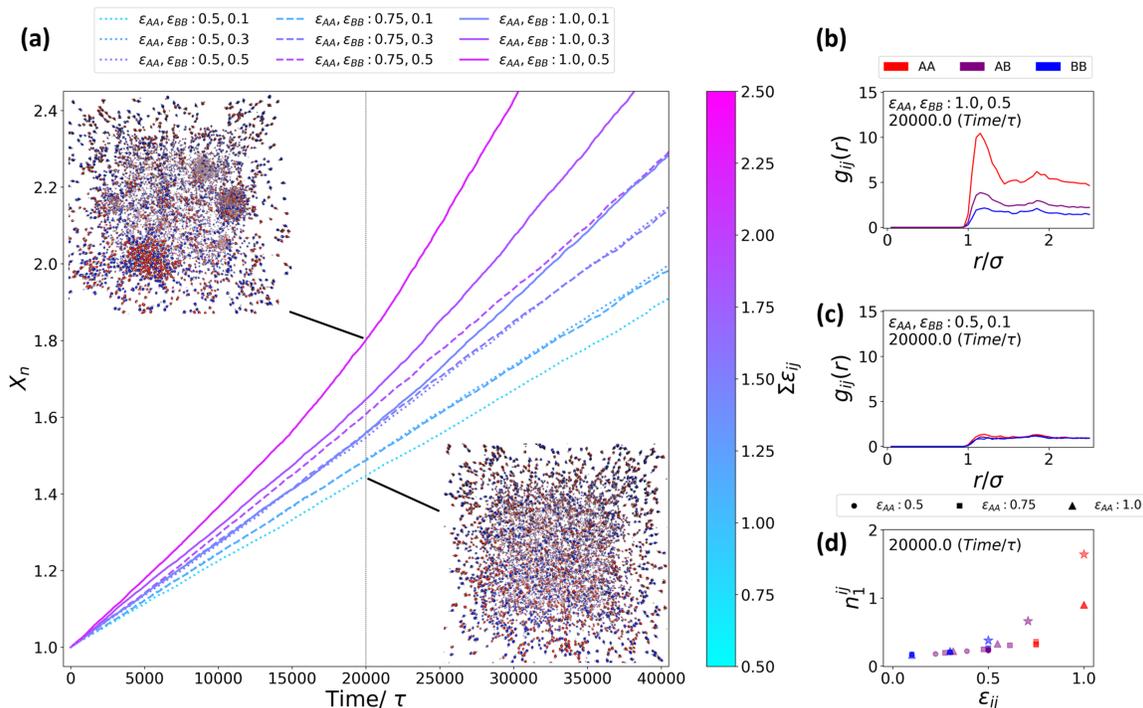


Figure 4.1: **Local monomer density and polymerization kinetics.** (a) Number averaged degree of polymerization  $X_n$  vs. simulation time for each combination of attractive interaction strengths explored. Dotted lines indicate simulations with  $\epsilon_{AA} = 0.5k_B T$ , dashed lines indicate  $\epsilon_{AA} = 0.75k_B T$ , and solid lines indicate simulations with  $\epsilon_{AA} = k_B T$ , while the coloration indicates the total attraction strength  $\Sigma \epsilon_{ij}$  for the simulation. The vertical dotted line shows a reference time of interest  $t = 20,000 \tau$ , at which the calculations in (b-c) and the inset structures were determined. Inset: A representative system structure for simulations with  $\epsilon_{AA} = k_B T$ ,  $\epsilon_{BB} = 0.5k_B T$ , and with  $\epsilon_{AA} = 0.5k_B T$ ,  $\epsilon_{BB} = 0.1k_B T$  as indicated, taken at the reference time  $t = 20,000 \tau$ . Type **A** monomers are shown in red, and type **B** monomers in blue. (b-c) Radial distribution functions  $g_{ij}(r)$  for (b)  $\epsilon_{AA} = k_B T$ ,  $\epsilon_{BB} = 0.5k_B T$  and (c)  $\epsilon_{AA} = 0.5k_B T$ ,  $\epsilon_{BB} = 0.1k_B T$ . The  $g_{ij}(r)$  function was calculated by monomer species pair and excluded nearest bonded neighbors. Indices  $i$  and  $j$  indicate the monomer species pair considered and are colored red for  $ij = \mathbf{AA}$ , blue for  $ij = \mathbf{BB}$ , and purple for  $ij = \mathbf{AB}$  or  $ij = \mathbf{BA}$ . (d) First coordination number,  $n_1^{ij}$ , vs.  $\epsilon_{ij}$  is shown by monomer type pairs for each attraction strength explored.

progresses approximately linearly at this new rate for the remainder of the reaction.

The mechanism behind this emergent rate change can be seen in the configuration of

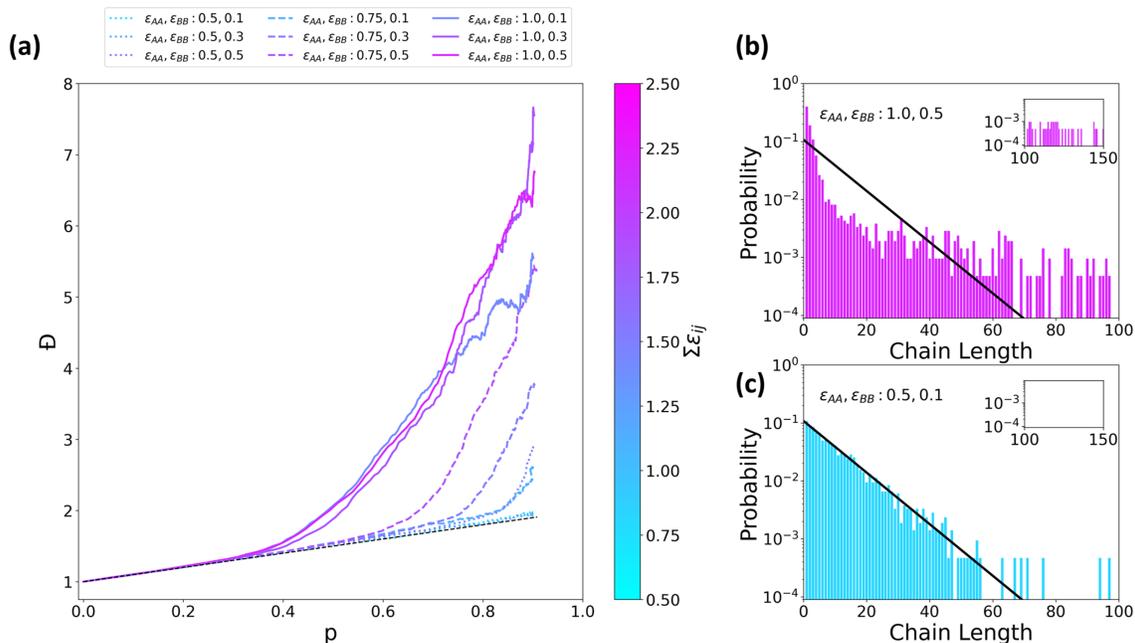


Figure 4.2: **Chain length distribution and dispersity.** (a) System dispersity,  $\mathcal{D}$ , as a function of reaction extent,  $p$ , for the full range of attractive interaction strengths explored. Dotted lines indicate simulations with  $\epsilon_{AA} = 0.5k_B T$ , dashed lines indicate  $\epsilon_{AA} = 0.75k_B T$ , and solid lines indicate simulations with  $\epsilon_{AA} = k_B T$ . The coloration indicates the total attraction strength  $\Sigma \epsilon_{ij}$  for the simulation, where  $ij$  are indices which correspond to each possible monomer species pair, namely **AA**, **AB**, and **BB**. The dashed black line shows the predicted  $\mathcal{D}$  from Flory-Schulz statistics of  $\mathcal{D} = 1 + p$ . (b-c) The chain length distribution at  $p = 0.9$  for (b)  $\epsilon_{AA} = 0.5k_B T$ ,  $\epsilon_{BB} = 0.1k_B T$  and (c)  $\epsilon_{AA} = k_B T$ ,  $\epsilon_{BB} = 0.5k_B T$ . The solid black lines show the expected Flory-Schulz distribution for  $p = 0.9$ . All results were averaged over three simulation trials for each parameter set.

the solution environment during the transition, as quantified by the radial distribution function between different monomer species,  $g_{ij}(r)$ , shown in Fig. 4.1b&c. Here  $ij$  is an index which correspond to the possible monomer type pairs within the system, namely **AA**, **AB**, or **BB**. For the cases in which attractive interactions are sufficient to drive the observed rate transition, there is a noticeable enrichment in the  $g_{ij}(r)$  for the associated monomer type pairs at the time when the rate change occurs – see Fig. 4.1b. By contrast, such an enrichment in local monomer type concentration is not present in the cases with lower overall inter-monomer attractions – no significant spatial correlation between the monomer type

pairs is observed in the  $g_{ij}(r)$  in Fig. 4.1c. To further quantify this behavior, we calculated the first coordination number,  $n_1^{ij}$ , for each set of attraction strengths and monomer species pair  $ij$  by integrating  $g_{ij}(r)$  over the first coordination shell (see Sec. B.2 in Appendix B for calculation details). The values obtained for  $n_1^{ij}$ , taken at  $t = 20,000\tau$  where the rate transition occurs for the strongest attraction case, are presented as a function of the governing attraction strength,  $\epsilon_{ij}$ , in Fig. 4.1d. The results clearly demonstrate the relationship between the first coordination number and the associated attractive interactions for each monomer type pair. For lower attraction strengths, which have not yet undergone a rate transition and for which standard kinetics still hold, there is a small but robust positive correlation between  $n_1^{ij}$  and  $\epsilon_{ij}$ . However, at the highest attraction strengths (star markers), which are in the midst of the rate transition at  $t = 20,000\tau$ , there is a sharp increase in  $n_1^{ij}$ , greater than that expected from the trend due to attraction strengths alone. This transition to a more aggregated phase is expected as a consequence of Flory-Huggins solution theory,<sup>73</sup> which explains that the entropic cost of forming a more condensed or demixed phase is reduced by the formation or lengthening of oligomers. The initial formation of oligomers shifts the balance of entropic and enthalpic contributions and favors the formation of an aggregated phase once chains are long enough so that the attractive interactions outweigh the entropic cost of demixing. The resulting increase in the local availability of reacting species in the aggregated phase accelerates the reaction, producing the change in rate behavior observed.

The distribution of chain lengths within the system is also affected by the alteration of the local solution environment produced by attractive interactions of sufficient strength. The Flory-Schulz distribution gives the ideal distribution of chain-lengths for a step-growth polymerization that proceeds with a constant rate.<sup>73</sup> Flory-Schulz theory also predicts a linear relationship between the dispersity,  $\mathcal{D}$ , and the reaction extent,  $p$ , such that  $\mathcal{D} = 1 + p$  for ideal polymerizations that proceed at a constant rate.<sup>73</sup> Figure 4.2a shows the observed  $\mathcal{D}$  as a function of the reaction extent,  $p$ . For combinations of attraction strengths that showed a transition in the kinetics such as that seen in the solid lines in Fig. 4.1a, there is an associated shift away from the linear Flory-Schulz prediction (dashed black line) to a non-linear  $\mathcal{D}$  behavior. Although not all transitions in rate occur within the time-window shown in Fig. 4.1a, we find that the rate increase coincides with the onset of non-linear oligomer dispersity in all such cases. The progression of  $\mathcal{D}$  and  $X_n$  with time and reaction extent over the full polymerization are available in Figures B.5 and B.6 in Appendix B.

As with the kinetic transition, the magnitude of the shift of this chain-length distribution away from ideal behavior depends on the magnitude of the attractions in the system,

with minimal attractions leading to a good agreement with the Flory-Schulz prediction, while progressively higher attractions show increasingly non-linear  $\mathcal{D}$ . The chain length distributions at  $p = 0.9$  for the highest and lowest total attraction strengths explored in this work are shown in Fig. 4.2b&c. For the lowest attraction strength in Fig. 4.2c, the expected Flory-Schulz distribution (solid black line) matches the observed chain length distribution. However, for the higher attraction strength shown in Fig. 4.2b, the observed chain length distribution is noticeably shifted towards longer chain lengths. Overall, Fig. 4.2 clearly demonstrates the breakdown in Flory-Schulz predictions caused by increased inter-monomer attraction, which broadens the chain length distribution and results in non-linear growth in  $\mathcal{D}$ .

### 4.3 Oligomer sequences depend on solvent selectivity

We have shown above how solvent-mediated asymmetric attractions between two monomer species can elicit collective behaviors that drive assembly, yielding non-standard kinetics and oligomer dispersities. We now examine the features of copolymer sequence that are directly impacted by this emergent aggregation process in order to probe if and how it alters the statistics of nearest neighbor monomers and yields long range correlations in oligomer sequences. Further, we examine how these sequence effects change as the difference in the solvent-mediated self-attractions between the two monomer species grows.

The direct sequence impact of the emergent co-localization of reacting species can be seen in the observed probabilities of sequential neighbors in the resultant oligomers. We denote these  $p_{AA}$  and  $p_{BB}$  for the probability of finding an **AA** or **BB** sequence, respectively, or  $p_{AB}$  for the probability of finding an **AB** or **BA** sequence. The analysis of the sequential neighbor probabilities, as determined at reaction extent  $p = 0.9$ , is shown in Fig. 4.3a-c for each of the attraction strengths explored in this study. Values are presented in terms of the deviation,  $\Delta p_{ij}$ , from the expected value for a truly random sequence,  $p_{ij}^{\text{random}}$ , i.e.,  $\Delta p_{ij} \equiv p_{ij} - p_{ij}^{\text{random}}$  for monomer type indices  $i$  and  $j$ . It should be noted that the values presented in Fig. 4.3 for  $p_{ij}$  change throughout the reaction as the sequences develop, as can be seen in Figures B.5 and B.6 in Appendix B.

For lower overall attraction strengths, where  $\varepsilon_{AA} = 0.5k_B T$  (bottom row), the sequences show minimal deviations from random behavior, such that  $\Delta p_{ij} \approx 0$  and there is a nearly equal probability of finding **AA**, **AB**, **BA**, or **BB** pairs within the resulting sequences. Where the difference between **A** to **A** and **B** to **B** attractions,  $\Delta\varepsilon \equiv \varepsilon_{AA} - \varepsilon_{BB}$ , increases

there is only a slight,  $\sim 2\%$  biasing towards **AA** pairs (bottom left corner). For the symmetric attraction case, where  $\epsilon_{AA} = \epsilon_{AB} = \epsilon_{BB} = 0.5k_B T$  and  $\Delta\epsilon = 0$ , the deviation from random sequences is negligible ( $\sim 10^{-3}$ ) for each sequence pair (bottom right corners). These random sequences are expected in the fully symmetric case, since the monomer species behave identically, differing only in their identifying label as **A** or **B**. For higher attraction strengths, a distinct trend emerges – as the attraction asymmetry increases, the sequence biasing of neighboring monomers also increases. For cases with  $\epsilon_{AA} > 0.5k_B T$  (top two rows), where the emergent kinetic shift and concentration enrichment occurs, sequences deviate markedly from random as biasing becomes much more significant. At the highest  $\epsilon_{AA}$  and  $\Delta\epsilon$  explored in this study (top left corner), **AA** pairs become the most likely sequence combination in the system, representing a  $\sim 15\%$  biasing away from random sequence behavior. Thus the strength and the relative balance of attractions between comonomer pairs are both integral to determining the resulting nearest neighbor probabilities.

The copolymer sequences obtained show evidence of long-range ordering as well, indicating that the sequence biasing effects extend beyond the nearest sequential neighbors. To explore this behavior, we make use of a metric developed for quantifying the extent of long range sequence correlations in DNA, which we adapted to the two-component copolymerization studied here.<sup>58,59,111</sup> This approach considers the sequence as a 1D random walk, with steps of +1 or -1 corresponding to **A** or **B** monomers, respectively. The metric,  $F(l)$ , then calculates the root-mean-square fluctuation of the sequence walk as a function of the distance  $l$  along the sequence. For a completely random sequence,  $F(l) \sim \sqrt{l}$ , and deviations from this scaling reflect a biasing in sequence over the length scale,  $l$ , where the deviation is observed. Fig. 4.3d shows  $F(l)$  vs.  $l$  for each set of attractive interactions explored. For cases with minimal difference in monomer self-attractions, for which  $\Delta\epsilon \leq 0.2k_B T$ , we observe no signs of longer-range sequence fluctuations. Here  $F(l)$  closely follows the scaling behavior predicted for a random sequence, which is expected given the minimal nearest neighbor sequence biasing seen in these combinations. However, as  $\Delta\epsilon$  increases, long-range sequence correlations appear, as indicated by the increased slope of  $F(l)$  for these cases. These correlations persist over a length scale greater than  $X_n \approx 10$ , the average chain length within the system.

The result of these long-range correlations in sequence may be seen in the length distributions of contiguous blocks of either **A** or **B** monomers within the oligomers. Fig. 4.3e-g shows the block length distributions for a fixed  $\epsilon_{AA} = k_B T$  and an increasing  $\Delta\epsilon$ . Similar to the Flory-Schulz distribution for chain lengths, both the block length distribution and

the dispersity expected from Markovian statistics can be calculated. In this calculation, the identity of each monomer only depends on the preceding monomer in the sequence and the observed nearest-neighbor bonding probabilities for that system –  $p_{AA}$ ,  $p_{AB}$ , and  $p_{BB}$  (see Chapter 2). The predicted Markovian distributions are then plotted along with the observed block length distributions for each monomer type. As  $\Delta\epsilon$  increases, the block distribution of the more self-attractive **A** monomer shifts to noticeably greater lengths than predicted by Markov statistics, displaying a broadening in the distribution analogous to that seen in the overall chain lengths at higher  $\Sigma\epsilon_{ij}$  values (see Fig. 4.2b). Interestingly, these shifts in block lengths are isolated to the **A** blocks, as the **B** block distributions are still well described by Markovian statistics. At larger  $\Delta\epsilon$ , the **A** monomers can be sufficiently attractive to condense into an **A**-enriched phase as polymerization proceeds and the oligomer chains lengthen, while the less attractive **B** monomers incorporate less into oligomers (see Fig. S4) and then less readily aggregate once in the oligomer phase. This incorporation bias results in the long-range correlations we observe in Fig. 4.3d and causes a breakdown in Markovian statistics, which promotes the formation of longer **A** blocks than can be explained by a Markovian model alone. In such cases, the self organization of reactants produces “blockier” oligomers, with block regions ordered along the oligomer chain by the relative self-attractions of their respective monomer species and the influence of the local environment. The ordering of block regions results in separate domains rich in **A** blocks and **B** blocks, analogous to the biased, but mixed, sequences found in gradient copolymers.

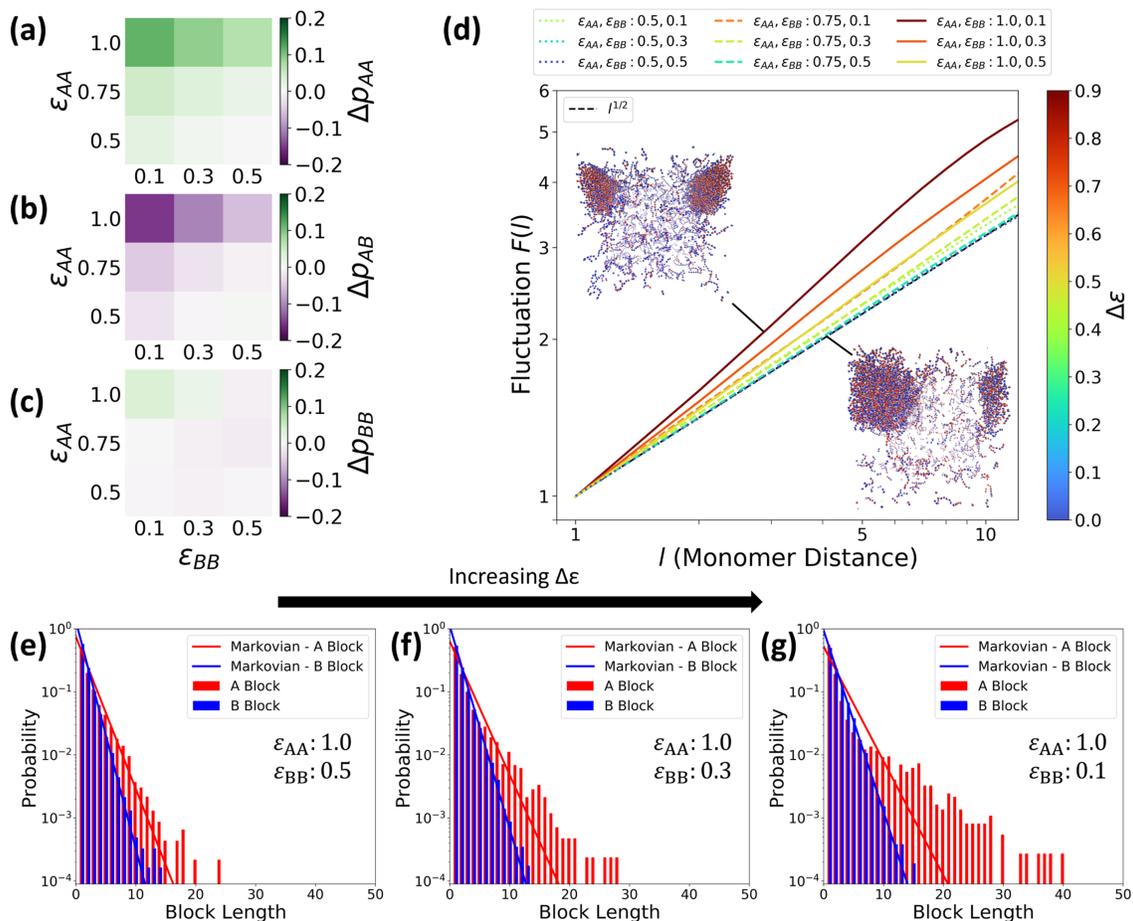


Figure 4.3: **Nearest neighbor and long-range sequence statistics.** (a-c) Heatmaps showing nearest neighbor bonding probabilities for (a) **AA**, (b) **AB**, and (c) **BB** pairs, for the full set of attractive interactions explored. Values shown are the differences,  $\Delta p_{ij}$ , between the observed probabilities and the probabilities expected for random sequences,  $p_{ij}^{\text{random}}$ , where indices  $i$  and  $j$  specify the monomer species pair. (d) Sequence fluctuations metric,  $F(l)$ , for all attraction strength combinations. The dotted black line shows the  $l^{1/2}$  scaling expected for a random sequence. Other dotted lines indicate simulations with  $\epsilon_{AA} = 0.5k_B T$ , dashed lines indicate  $\epsilon_{AA} = 0.75k_B T$ , and solid lines indicate simulations with  $\epsilon_{AA} = k_B T$ , while the coloration indicates the  $\Delta\epsilon$  value for the simulation set. Inset: Representative system structures for systems with  $\Delta\epsilon = 0$  and  $\Delta\epsilon = 0.9k_B T$ . Type **A** monomers are shown in red, and type **B** monomers are shown in blue. (e-g) Block length distributions for systems with  $\epsilon_{AA} = k_B T$  and (e)  $\epsilon_{BB} = 0.5k_B T$ , (f)  $\epsilon_{BB} = 0.3k_B T$ , and (g)  $\epsilon_{BB} = 0.1k_B T$ . The expected distribution for Markovian statistics for the **A** and **B** blocks are plotted as red and blue lines respectively, on top of the observed distribution of block lengths of each type as histograms in the same color. Analyzed sequences and block length distributions were collected from three independent trials for each parameter set, all at reaction extent  $p = 0.9$ .

## 4.4 Oligomer persistence length impacts both sequence and aggregate structure

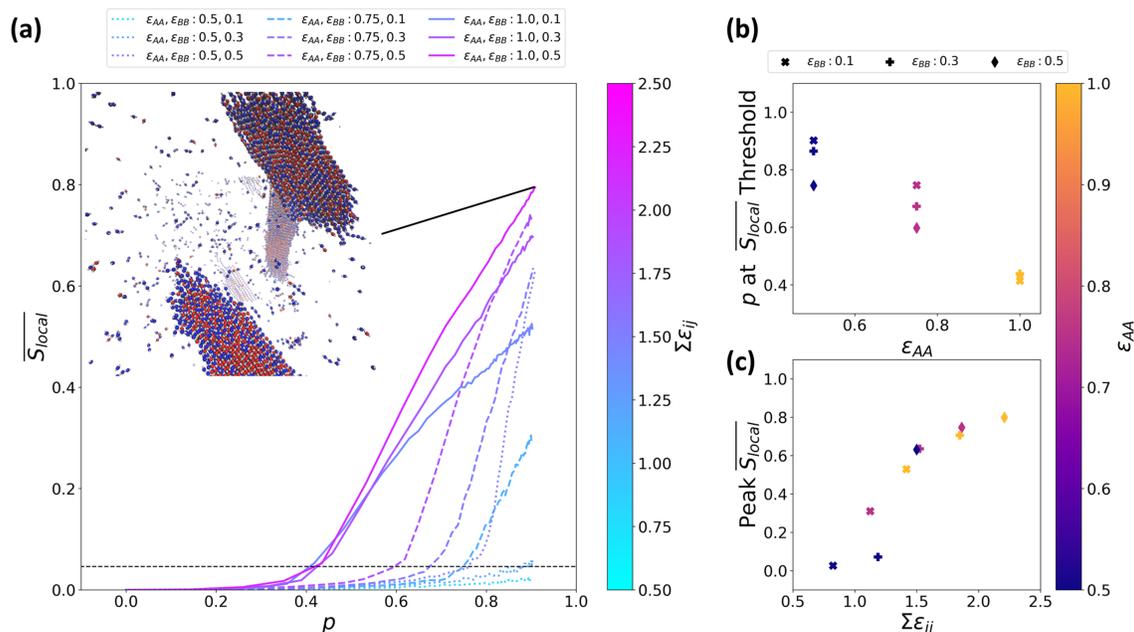


Figure 4.4: **Localized chain alignments in stiff-chain oligomers.** (a) The local nematic ordering parameter,  $\overline{S}_{\text{local}}$ , is plotted here as a function of reaction extent,  $p$ , over the range of attraction strengths explored for stiff-chain,  $l_p = 16.5$ , polymers. Dotted lines indicate simulations with  $\epsilon_{AA} = 0.5k_B T$ , dashed lines indicate  $\epsilon_{AA} = 0.75k_B T$ , and solid lines indicate  $\epsilon_{AA} = k_B T$ , while the coloration indicates the total attraction strength,  $\sum \epsilon_{ij}$ , for the simulation. The dashed horizontal black line indicates the highest value for  $\overline{S}_{\text{local}}$  seen for the flexible-chain,  $l_p = 3.5$ , polymers (see Fig. S3), and is used as a threshold to determine early emergence of chain alignment. Inset: A representative system structure at  $p = 0.9$  for simulations with  $\epsilon_{AA} = k_B T$ ,  $\epsilon_{BB} = 0.5k_B T$ . Type **A** monomers are shown in red, and type **B** monomers in blue. (b) The reaction extent at which  $\overline{S}_{\text{local}}$  exceeds the threshold value is shown for each parameter set as a function of  $\epsilon_{AA}$ . (c) The peak value of  $\overline{S}_{\text{local}}$  reached for each simulation parameter set as a function of  $\sum \epsilon_{ij}$ . For both (b) and (c), coloration indicates  $\epsilon_{AA}$  value and marker style indicates  $\epsilon_{BB}$  value. Results were averaged over three simulation trials for each parameter set.

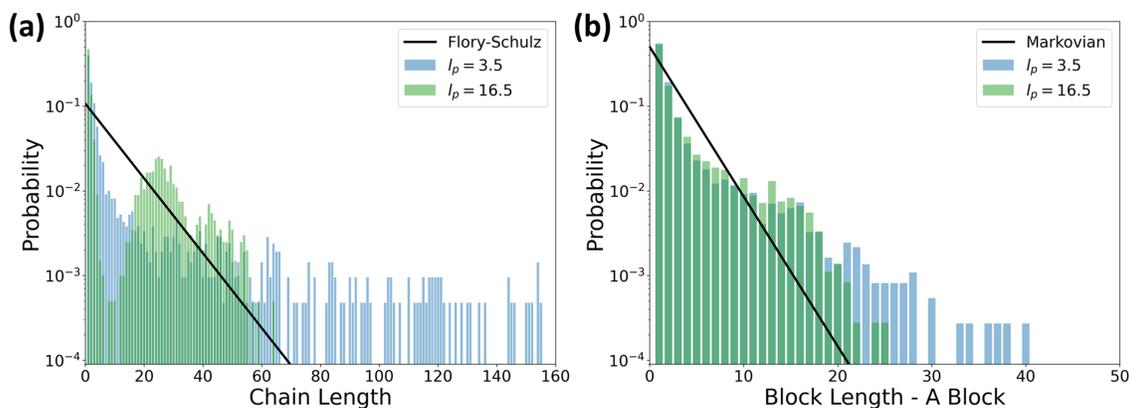
Previously, we have observed the formation of pseudo-crystalline nematic alignment in oligomers with sufficiently high persistence lengths and symmetric **A** to **A** and **B** to **B**

attractive interactions.<sup>59</sup> This alignment impacted both the dispersities and the oligomer chain and block length distributions within the system, in a manner dependent on oligomer stiffness. We now discuss this emergent transition to nematic alignment and examine its impact on oligomer chain and block lengths in stiff oligomers under the conditions of asymmetric attractive interactions.

In order to quantify the formation of these aligned structures, we use an order parameter,  $\overline{S}_{\text{local}}$ , which is an ensemble averaged measure of the local orientational ordering within regions of sufficient density (see Sec. B.3 in Appendix B for details). Like other common ordering parameters used in the study of nematic phases in liquid crystals,  $\overline{S}_{\text{local}}$  transitions from  $\overline{S}_{\text{local}} = 0$  for an isotropic system to  $\overline{S}_{\text{local}} = 1$  for a perfectly aligned system. The introduction of a density criteria and the restriction to measuring local alignment focuses this parameter on alignment within individual oligomer aggregates, allowing for high  $\overline{S}_{\text{local}}$  values in systems with internally aligned aggregates, even if independent aggregates do not share the same orientation.  $\overline{S}_{\text{local}}$  therefore captures the extent of local alignment among nascent oligomers but does not describe a system-wide preferred orientation as in a true liquid crystal nematic phase.

Fig. 4.4a shows the progression of  $\overline{S}_{\text{local}}$  as a function of the reaction extent,  $p$ , for stiff chain oligomers of  $l_p = 16.5$  at each of the attractive interaction strengths explored in this work. Orientationally ordered aggregates are clearly observed for  $\Sigma\varepsilon_{ij} \gtrsim 1.2k_B T$ , where  $\overline{S}_{\text{local}}$  reaches values  $\geq 0.5$ , indicating significant alignment amongst neighboring chains. For the highest value of  $\Sigma\varepsilon_{ij}$  at  $p = 0.9$ , shown in the inset of Fig. 4.4a, nearly all of the condensed oligomers are orientationally-ordered. The observed nematic alignment is sensitive to both polymer attraction strengths and chain stiffness. For flexible polymers,  $\overline{S}_{\text{local}}$  values remain close to zero (see Fig. B.3 in Appendix B), and even higher attractions display only a marginal increase in  $\overline{S}_{\text{local}}$ . By contrast, at all but the lowest attraction strengths, stiff polymer alignment surpasses the highest alignment seen in any flexible chain system, a threshold value indicated by the dashed black line in Fig. 4.4a. Though higher attraction strengths promote the formation of the condensed aggregates, it is the higher persistence length that restricts the accessible conformations of chains in those aggregates, leading to their orientational ordering. A notable feature of the ordering behavior is that the onset of nematic alignment and the final peak value for  $\overline{S}_{\text{local}}$  are sensitive to different aspects of the attractive interactions. Fig. 4.4b&c show the trend in these properties as a function of  $\varepsilon_{AA}$  and  $\Sigma\varepsilon_{ij}$  respectively. Fig. 4.4b shows that the point at which  $\overline{S}_{\text{local}}$  exceeds the threshold established by the flexible chain case is predominantly controlled by  $\varepsilon_{AA}$ . For  $\varepsilon_{AA} = k_B T$  in particular, this onset of ordering is largely independent of the value of  $\varepsilon_{BB}$  and thus

$\Delta\epsilon$ . In this case, when  $\epsilon_{BB}$  is small, **A** to **A** attractions are sufficient to drive alignment of **A**-enriched oligomers, which form earlier in the reaction when the incorporation of **B** monomers into the polymer phase is significantly less than that of **A** (see Fig. S3). The  $\overline{S}_{\text{local}}$  values obtained from early alignment are surpassed later in the reaction, however, by cases with higher total attractions,  $\Sigma\epsilon_{ij}$ , once more **B** monomers have been incorporated into the oligomers. The  $\overline{S}_{\text{local}}$  values at  $p = 0.9$  are thus predominantly controlled by  $\Sigma\epsilon_{ij}$  as seen in Fig. 4.4c. The variations in these two stages of nematic ordering further demonstrate how differences in non-bonded attractive interactions can alter the formation of the aggregate phase as the reaction proceeds.



**Figure 4.5: Chain and block length distributions and their variation with persistence length.** (a) Chain length distributions are shown here at  $p = 0.9$  for systems with  $\epsilon_{AA} = k_B T$ ,  $\epsilon_{BB} = 0.5k_B T$ . Distributions of both flexible chains of  $l_p = 3.5$  (blue) and stiff chains of  $l_p = 16.5$  (green) are plotted. The solid black line shows the expected Flory-Schulz distribution for  $p = 0.9$ . (b) Block length distributions for contiguous **A** monomer blocks are plotted at  $p = 0.9$  for systems with  $\epsilon_{AA} = k_B T$ ,  $\epsilon_{BB} = 0.1k_B T$  for both flexible chains ( $l_p = 3.5$ , blue) and stiff chains ( $l_p = 16.5$ , green). The solid black line shows the expected distribution from Markovian statistics. (NB:  $p_{AA}$  is approximately the same for both stiff and flexible chains at these attraction strengths, so the associated Markovian prediction is equivalent.) Distributions in (a) and (b) were collected from three independent trials for each persistence length.

The impact of chain stiffness and nematic alignment on chain and block lengths is significant. Fig. 4.5 shows the chain and block length distributions at  $p = 0.9$  for systems with two different persistence lengths but the same set of attractive interactions. In both cases, attractive interactions strong enough to promote aggregation show a deviation from Flory-

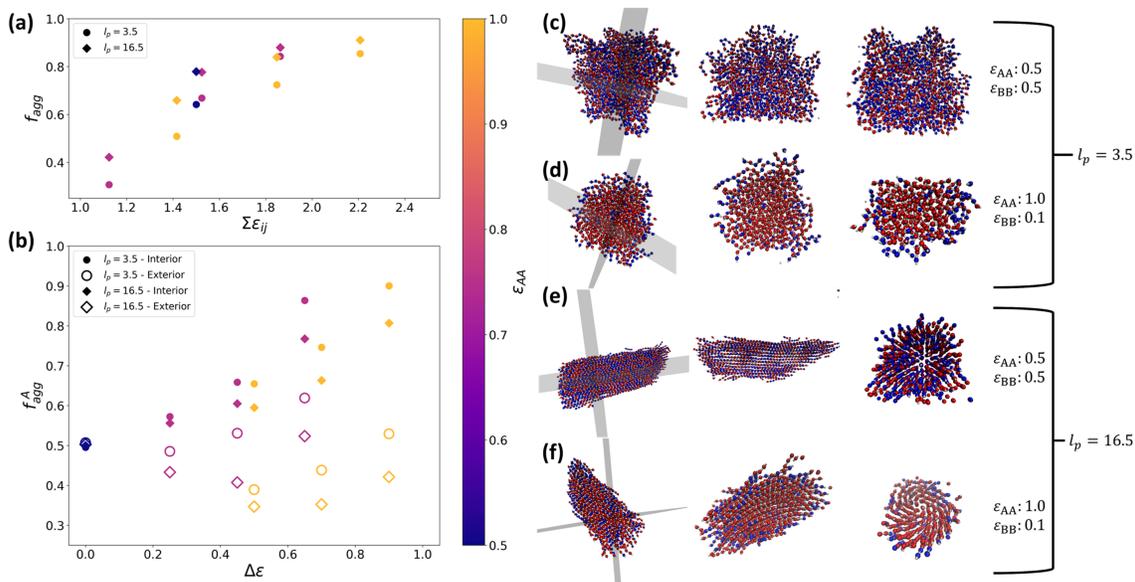
Schulz and Markovian statistics, but the nature of this deviation differs significantly. As previously discussed, flexible chains show an increased dispersity and a broadened block length distribution, due to the general promotion of longer chain and block lengths in the system. By contrast, stiff chains show a truncation in both chain and block length distributions and an enrichment at specific chain lengths, indicative of a characteristic length introduced to the system by the persistence length and associated orientational ordering. These results are similar to what was previously seen in the case of symmetric attractions.<sup>59</sup> However, with the asymmetric attractions explored in this work, the non-Markovian sequence behavior is isolated to the **A** block lengths (Fig. S5& S6), which can also be seen in the flexible chain case (Fig. 4.3). These results illustrate the complex interplay between the attractive interactions and the chain stiffness. This interplay governs the phase behavior and accessible conformations of oligomers and, in turn, influences the sequence of the resulting chains as well as the morphology of their aggregates.

## 4.5 Nascent sequence and selective solvation influence the composition of different regions within the aggregates

The alterations to the ideal reaction kinetics and sequences described above are associated with the formation of a condensed, aggregated phase driven by oligomer formation. Here we explore the composition and structure of these oligomer aggregates.

Nascent chain aggregation, the emergence of which was described in Fig. 4.1 for the early stages of the reaction, continues throughout the reaction with the continued growth of oligomer chains, an expected consequence of Flory-Huggins theory.<sup>73</sup> After 90% of possible bonds in the simulation have formed, this aggregated phase has developed into clearly distinct structures of clustered oligomers. The nematic alignment observed at high persistence lengths is fundamentally driven by the formation of dense oligomer aggregates. To identify properties of this aggregate phase in the late stage of the reaction, we make use of a local neighbor metric,  $n^{\text{local}}$ , which is defined as the number of neighboring monomers that are within a distance of  $2.5\sigma$ . Threshold values for this metric were chosen both to distinguish aggregated structures from the surrounding dilute phase and to delineate regions within these structures that broadly correspond to their exteriors and interiors. The  $n^{\text{local}}$  metric also reports on the relative solvent accessibility of each region within the aggregate. We set two threshold values for  $n^{\text{local}}$  at 12 and 36 neighboring monomers. The interior of the aggregates are then defined by monomers with  $n^{\text{local}} \geq 36$ , while the exterior of the

aggregates are defined by monomers with  $12 \leq n^{\text{local}} < 36$ . Monomers with  $n^{\text{local}} < 12$  are not considered to be part of the aggregated phase. According to these criteria, the densities of the exterior regions of an aggregate are at least three times the initial bulk density for monomers in the simulation box, while the densities of the interior regions are at least nine times the initial bulk density.



**Figure 4.6: Aggregate compositions and representative structures.** (a)  $f_{\text{agg}}$ , the fraction of all monomers in the system that are located within an aggregate (as defined by  $n^{\text{local}} \geq 12$ ), is shown as a function of the total attraction strength,  $\Sigma \epsilon_{ij}$ . (b)  $f_{\text{agg}}^A$ , the fraction of **A** monomers within an aggregate, is plotted as a function of  $\Delta \epsilon$ . Monomers with  $n^{\text{local}} \geq 36$  were identified as “interior” (filled markers), and monomers with  $12 \leq n^{\text{local}} < 36$  were identified as “exterior” (hollow markers). For both (a) and (b), persistence length is indicated by marker shape while coloration indicates the value of  $\epsilon_{AA}$ . Aggregate populations were collected from three independent simulation trials per parameter set. (c-f) Representative aggregate structures identified by the neighboring criteria employed in (a&b) for simulations with (c)  $\Delta \epsilon = 0$ ,  $l_p = 3.5$ , (d)  $\Delta \epsilon = 0.9$ ,  $l_p = 3.5$ , (e)  $\Delta \epsilon = 0$ ,  $l_p = 16.5$ , and (f)  $\Delta \epsilon = 0.9$ ,  $l_p = 16.5$ . The first column shows a single aggregate structure along with orthogonal cross-section planes. The second and third columns show the interior of each aggregate structure, as viewed from each of the shaded cross-sections. Type **A** monomers are shown in red, with type **B** monomers in blue. All results are shown at  $p = 0.9$ .

Fig. 4.6 shows the results of the application of these local neighbor criteria to our system

at  $p = 0.9$ . Notably, for systems with  $\epsilon_{AA} = 0.5k_B T$  and  $\epsilon_{BB} < 0.5k_B T$ , we do not find regions for which  $n^{\text{local}} \geq 12$ . At these attractions, the reaction kinetics (Fig. 4.1a) and dispersities (Fig. 4.2) are as expected for an ideal step-growth process, the  $g(r)$  function shows no concentration enrichment early in the reaction (see Fig. 4.1c), and no aggregates are visible in the system at  $p = 0.9$ . The correspondence among these results suggests that the  $n^{\text{local}}$  criterion is appropriately delineating the aggregated phase. In Fig. 4.6a, the total fraction of all monomers in the aggregated phase,  $f_{\text{agg}}$ , is shown as a function of the total attraction strength,  $\Sigma\epsilon_{ij}$ . As with the aggregation and the associated rate change that occurs earlier in the polymerization (see Fig. 4.1), we find that the total incorporation into aggregates at the late stage of the reaction predominantly depends on the total strength of the intermonomer attractions.

Having established criteria for identifying the interior and exterior regions of the aggregates, we next sought to quantify the incorporation of each monomer species into these regions, examining the fraction of **A** monomers within the aggregate,  $f_{\text{agg}}^{\text{A}}$ , within each region. The results for  $f_{\text{agg}}^{\text{A}}$  are shown in Fig. 4.6b for all simulation parameters where aggregates form. For the symmetric attraction case,  $\Delta\epsilon = 0$ , both monomer species are equally incorporated into the interior and exterior regions of the aggregates. Representative aggregate structures under symmetric attractions (Fig. 4.6c&e) clearly show the even incorporation of both monomer species throughout the entire aggregate. At the onset of attraction strength asymmetry, however, a distinct enrichment of **A** monomers in the interior regions is apparent. For modest attraction asymmetries, this enrichment of **A** in the interior is associated with a depletion of **A** in the exterior. As the attraction asymmetries increase further, however, this depletion of **A** in the exterior is actually lessened, because fewer **B** monomers are incorporated into the aggregates, remaining instead in the surrounding dilute phase (see Fig. S4). In the maximally asymmetric case, **B** monomers are significantly less incorporated and almost entirely located in the aggregate's exterior, as can be seen in the representative structures in Fig. 4.6d&f.

The impact of increased chain stiffness on the formation and structure of aggregates is also apparent from the results in Fig. 4.6. Increased chain stiffness has been shown to influence the phase behavior of polymers, and greater stiffness further promotes the formation of an aggregated phase.<sup>116–120</sup> This behavior is borne out in our system, as stiff chain oligomers show greater incorporation into the aggregated phase than flexible oligomers for all attraction strength combinations explored (Fig. 4.6a). Further, the interiors of stiff oligomer aggregates are less enriched in **A** than in the corresponding flexible oligomer aggregates with the same inter-monomer attractions – see Fig. 4.6b. This reduction of interior

$f_{\text{agg}}^{\text{A}}$  at increased persistence length is a consequence of two features of the increased chain stiffness. First, the increased total aggregation which stiff chains promote (Fig. 4.6a) necessarily increases the incorporation of **B** throughout aggregates. Secondly, greater chain stiffness restricts the accessible conformations of the oligomers within the aggregate. Flexible chains may readily adopt conformations that position incorporated **B** monomers in the exterior of an aggregate, even when an isolated **B** monomer is located within a stretch of **A** monomers in the chain. In stiff chains, however, such bent conformations are restricted, and **B** monomers incorporated within the interior of a chain are likely to be located in the interior of the aggregate as well. Such conformational restrictions also control the morphology of the aggregate in a manner largely independent of the attraction strengths, with flexible chains forming loosely spherical, globule structures (Fig. 4.6c&d), and stiff chains forming elongated, rod-like aggregates (Fig. 4.6e&f).

It is important to note that the extent to which each monomer species is incorporated into the growing oligomers changes throughout the reaction in a manner dependent on the balance of attraction strengths between the monomer species. For high attraction strength asymmetry, we observe an enrichment in the more attractive **A** monomers earlier in the reaction and an increased incorporation of **B** later on. Additional data on the polymerized fraction of each monomer species and its variation with  $p$  is shown in Fig. S4.

## 4.6 Summary and conclusions

In this Chapter, we have explored the influence of differing solvent affinities in step-growth copolymerization by modulating the effective inter-monomer attractions acting between each monomer combination. We find that, in cases with even mild solvent-mediated attractions, an emergent co-localization of reacting species occurs which promotes the formation of an aggregate phase as the reaction proceeds. This spatial heterogeneity produces non-standard kinetic effects, changing both the reaction rate and the resulting dispersity. The self-assembly of the reacting species also influences the sequence of the oligomers formed, in a manner dependent on the selectivity of the solvent interactions. The complexities arising from these collective behaviors lead to associated kinetic and sequence effects that cannot be fully captured by standard Flory-Schulz or Markovian statistical descriptions.

Understanding the impact of solvent affinities on sequence development and phase behaviors in solution-based copolymerizations is an important step towards developing generalizable synthetic approaches to sequence controlled or sequence-biased polymers for

targeted supramolecular assemblies. A greater comprehension of the collective behavior, biased bond formation, and nascent chain self-assembly that can arise during the copolymerization of monomers with differing solvent affinities will improve our capacity to harness these effects towards the design and development of advanced copolymeric materials.

## Chapter 5

# Nematic Ordering and Emergent Characteristic Block Length

### 5.1 Introduction

In the preceding Chapters,<sup>86,87</sup> we have shown that differences in non-bonded attractive interactions that approach the strength of thermal fluctuations ( $\sim k_B T$ ), such as those arising from differences in monomer-solvent affinities, can significantly influence the sequences obtained from a solution-based step-growth polymerization. This sequence influence is due to the growth of nascent oligomers, which alters the free energy of mixing as explained by Flory-Huggins solution theory, leading to a microphase separation of reactants as the reaction proceeds.<sup>73</sup> Due to this local demixing, as oligomers lengthen, attractive interactions drive the formation of emergent concentration heterogeneities in the local reaction environment, altering the likelihood of sequence additions. The chain stiffness of the growing oligomers, as characterized by the persistence length, also influences both the sequence development and the self-assembly of the resulting aggregates, by altering the behavior of the emergent demixed microphases.<sup>59,86,87</sup> Oligomers with higher persistence length can form highly ordered, nematically aligned phases as the reaction proceeds, giving rise to

---

The work in this Chapter is has been adapted with minor alterations from a manuscript in development. Ryan L. Hamblin contributed to the investigation, software, formal analysis, visualization, and writing. Zhongmin Zhang contributed to the investigation, software, formal analysis, and visualization. Kateri H. DuBay contributed to the conceptualization, methodology, supervision, formal analysis, and editing.

what we term a self-templating effect that alters subsequent chain growth. Both the length of the chains formed and their “blockiness”, *i.e.*, the relative lengths of sequences of repeated monomer additions of a single type, are influenced by this effect. Such templating behaviors lead to the development of characteristic lengths, in which the reaction is biased towards the formation of particular chain and block lengths. In this Chapter, we investigate the emergence of these characteristic lengths in time, examining how the bonding behaviors of stiff-chained copolymers differ from those with more flexible chains leading to altered sequence ensembles. Further, we explore how features of the system which influence reaction kinetics and the relevant characteristic timescales within the system alter these bonding behaviors at different stages of the reaction, thereby impacting the chain and block length statistics. To this end, we vary the activation energy of the reaction, the solution viscosity, and the initial monomer density, to examine their effects on the kinetics of the reaction, the alignment of the growing oligomers, and the resulting chain and sequence statistics. We find that the characteristic length scale we observe shifts in a predictable fashion in response to the characteristic system timescales defined by diffusion and reaction kinetics in the system.

We begin, in Section 5.2, by examining how the sequence self-templating phenomenon we observe responds to changes in the reaction conditions, including the non-bonded interaction strength, reactant concentration, monomer reactivity, and solvent viscosity. In Section 5.3 we relate these reaction conditions to characteristic timescales of the system, noting the variation in the onset of sequence templating and the resulting characteristic length scale shifts predictably in response to changes to the rates of reaction and monomer diffusion. In Section 5.4 we describe how this self-templating sequence behavior emerges in time, arising as a consequence of the simultaneous impact of emergent demixing and nematic ordering, with an associated change in bonding behaviors in the system. In Section 5.5 we explore the nucleation and growth of oligomer aggregates, examining how aggregate morphology drives these changes to bonding behavior within the aggregated phase in a manner that depends on the length and stiffness of the reacting chains. We then briefly conclude in Section 5.6, discussing the potential for exploiting the observed phenomenon to influence the sequence development of stiff-chain copolymers in the context of simple, one-pot reaction schemes.

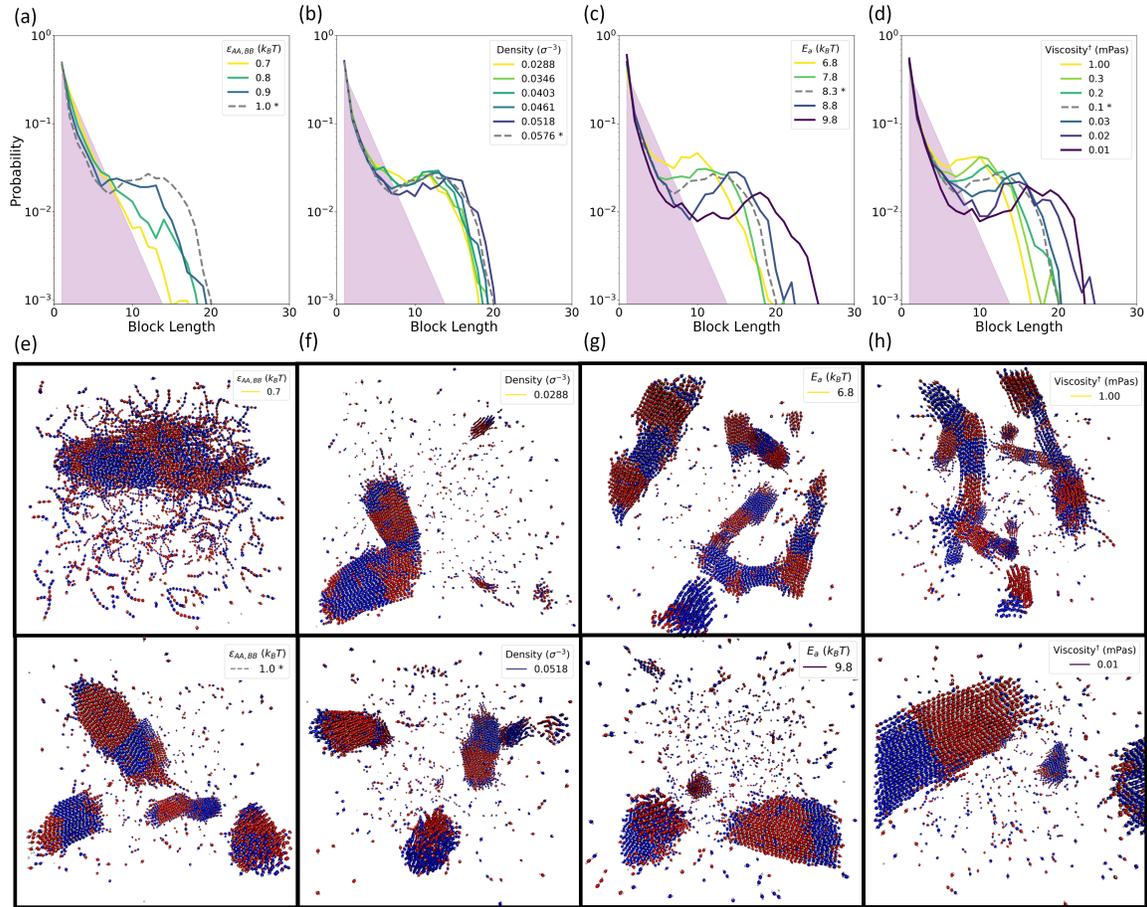


Figure 5.1: **Block length distributions vary with attraction strength, monomer density, and viscosity.** (a-d) Block length distributions of all-A or all-B blocks are shown for the copolymerization of stiff chains ( $l_p = 16.5$ ) at a reaction extent of  $p = 0.9$ . Results are shown here for variations in (a) like-monomer attraction strength, (b) initial monomer density, (c) activation energy, and (d) solvent viscosity. The gray, dashed lines in each plot are identical and display the block length distribution found under the standard conditions described in Ref. [59]. These conditions are indicated with \* in the legends, which also specify the values of the conditions held constant in (a-d). Each distribution is obtained from three independent simulation trials at a reaction extent  $p = 0.9$ . The shaded purple region corresponds to the results expected from Markov statistics for the case with  $\epsilon_{AA, BB} = 0.7 k_B T$  pictured in (e), where we observe the onset of polymerization driven assembly behaviors. (e-h) Visualization of representative structures under select conditions from (a-d). Each column corresponds to changes in the same property. The center row showcases conditions with the minimal block length distribution shift, while the bottom row shows conditions with maximum shift in the block length distribution. Specific parameters for the visualizations are indicated by the inset legend, which correspond to block length distributions in (a-d).

## 5.2 Reaction conditions shift characteristic block length

In our previous work, we showed how copolymer chains of sufficient stiffness and non-bonded attractions yield an emergent characteristic block length, which alters the distribution of block and chain lengths and reduces their dispersity.<sup>59</sup> Here we begin by considering the extent to which this sequence biasing phenomenon may be influenced by other properties of the reaction independent of the chain stiffness.

To investigate the dependence of the observed emergent characteristic all-A or all-B block length on factors other than the persistence length, we varied the initial bulk density of the monomers, the viscosity of the implicit solvent, the strength of non-bonded attractions, and the activation energy of bond formation. We therefore ran a series of simulations which modified one of each of these variables over a range of plausible values, while keeping the other variables fixed. For additional details on how these variable were adjusted in the simulation, refer to Sec. C.1.1 & C.1.2 in Appendix C. Fig. 5.1 shows the results of these simulations, examining the response of the block length distribution to changes in each of these parameters in Fig. 5.1(a-d) and showcasing sample structures for two conditions in each case in Fig. 5.1(e-h). In the plots shown in Fig. 5.1(a-d), the results for our set of standard conditions in previous studies<sup>58,59</sup> are the same across plots and indicated with a gray dashed line. Additional results showing the relative proportion of oligomer bonding, namely bonding involving oligomers of length  $\geq 3$  monomers, alongside block length distributions under the conditions shown in Fig. 5.1, are provided in Fig. C.8 in Appendix C.

The dependence of the emergent block length on each of these variables differs, and broadly falls into three categories. The first of these is the behavior of the system in response to a reduction in the strength of non-bonded attractions, namely, to a reduced value of  $\epsilon_{AA, BB}$ , which governs the strength of inter-monomer attractions. Here we see that the characteristic block length (Fig. 5.1a) rapidly disappears from the system with a reduction of the like-monomer attractions of only  $0.3k_B T$ . This result highlights the absolute dependence of the observed phenomenon on the emergent aggregation and assembly driven by effective non-bonded like-monomer attractions. The reduction in attraction strength modifies the phase behavior so that the cohesive energies are insufficient to overcome the entropic cost of aggregation until very late stages of the reaction when chains of sufficient length have formed (Fig. 5.1e, top). With the impact of the emergent aggregation and phase separation thus greatly reduced, the resulting block length distribution collapses towards the purple-shaded distribution expected from Markovian statistics, which describes the se-

quence behavior when the probability of each additional monomer in the chain depends only on the identity of the preceding monomer (Fig. 5.1a). By contrast, increased attractions result in aggregation and assembly occurring much earlier in the reaction, meaning a greater proportion of bond formation and sequence development occurs after oligomers begin to assemble. In such instances, the impact of collective behaviors on sequence is more robust, and the associated block distribution shift away from the Markovian distribution is more pronounced.

In contrast to the dramatic onset of new phase and sequence behaviors seen in response to small variations in attractive energies, the response to variation in the initial monomer density is minute. Even at a 50% reduction in the initial monomer density, the distribution of block lengths (Fig. 5.1b) remains largely unchanged, with the peaked distribution indicative of a characteristic block length remaining (Fig. 5.1f). It should be noted that polymerization induced phase separation as a whole depends upon reactant density, and we therefore necessarily restrict our exploration of density values to be low enough that reactants are initially well solvated, but large enough to undergo sufficient aggregation at oligomer lengths reached in the simulation as the reaction proceeds. Within this density regime, however, the aggregation behavior and emergence of the observed characteristic block length is not sensitive to the absolute rate of the reaction, a reactant density dependent property.

The third type of dependence we observe in Fig. 5.1 is in response to variations in the relative, rather than absolute, kinetic rates within the system, through changes to the activation energy of the reaction (Fig. 5.1c) and the viscosity of the implicit solvent (Fig. 5.1d). While a characteristic block length remains in evidence across variations in these timescales, changes to these variables shift its length. Changing these conditions also alters the timing and extent of oligomer bond formation, *i.e.*, the bonding that involves at least one chain of length  $\geq 3$  monomers in the system (see Fig. C.8 in Appendix C for oligomer bonding details). Not only do the block length distributions show similar shifts under variation to each of these quantities, but changes to the structure and morphology of the aggregates appear analogous as well. Conditions leading to a shorter characteristic length show smaller and more numerous aggregates (Fig. 5.1g&h, top), while single large aggregates dominate the systems in which conditions lead to a longer block length (Fig. 5.1g&h, bottom).

### 5.3 Block length distribution shifts correspond to changes in relative kinetic timescales

To better quantify the observed changes in characteristic block length biasing, we now consider the effective reaction and diffusion timescales in the system in terms of the activation energy of the reaction and the solvent viscosity.

In order to define a characteristic timescale for the reaction,  $\tau_R$ , we consider the inverse of the effective rate constant for polymerization in the absence of aggregation behaviors,  $k_{\text{eff}}$ , taking  $\tau_R \equiv k_{\text{eff}}^{-1}$ .<sup>121</sup> The kinetics for our model system have been worked out in detail in previous work,<sup>87</sup> allowing us to connect the measured activation energy,  $E_a$ , to our effective rate constant and thus our reactive timescale,  $\tau_R$ .

To characterize the impact of solvent viscosity on the diffusional timescale of the reactants, we make use of the Stokes-Einstein relation to determine a diffusion coefficient,  $D$ , in terms of the solvent viscosity,  $\eta$ , and the size of an individual monomer,  $\sigma$ . Specifically, we define the timescale of diffusional motion using  $\tau_D \equiv \sigma^2/D$ , i.e., the time needed for a monomer to diffuse its own length.

It should be noted that we expect the behaviors we observe to depend on the relative balance of these two timescales,  $\tau_D/\tau_R$ , and not on the values of the timescales themselves. As such, the trends we observe should not be influenced by constant factors that alternative definitions of  $\tau_D$  and  $\tau_R$  might introduce, so long as the functional forms of their dependence on the tunable parameters in our model,  $\tau_D(\eta)$  and  $\tau_R(E_a)$ , remain consistent.

With the chosen definitions of  $\tau_D$  and  $\tau_R$  in hand, we are able to map our variations in activation energy and solvent viscosity to the relative balance of diffusional and reactive timescales in our system. Fig. 5.2 shows the combined results of simulations from Fig. 5.1c&d, in which either solvent viscosity or reaction activation energy was varied, as quantified by the ratio of timescales,  $\tau_D/\tau_R$ , indicated by line coloration. The extent of the block distribution shift in Fig. 5.2a clearly corresponds to a shift in the ratio of these characteristic timescales,  $\tau_D/\tau_R$ , as can be seen in the color progression from left to right.

To quantify the observed shift in the block length distribution, we make use of the first-order Wasserstein distance,<sup>96,97</sup>  $W_1$ , as a measure of the statistical distance between our observed block length distribution and the block length distribution expected if sequence development obeyed Markovian statistics.<sup>58,86</sup> Utilizing this statistical distance,  $W_1$ , we find a direct quantitative relationship (Fig. 5.2a inset) between the degree of distribution shift and the ratio of characteristic timescales in the system. Alongside the clear correspon-

dence between block length distributions and coloration in the main panel of Fig. 5.2a, this quantitative relationship suggests that the effect of changes to both activation energy (Fig. 5.1c) and solvent viscosity (Fig. 5.1d) can be captured in a unified way through  $\tau_D/\tau_R$ .

The onset of nematic ordering, characterized by  $\overline{S_{\text{local}}}$  in Fig. 5.2b, is similarly impacted by changes to the balance of characteristic timescales – as are the chain and block length dispersities shown in Fig. 5.2c. Each of these features develop earlier or later in the reaction, and to a greater or lesser extent, at lower or higher  $\tau_D/\tau_R$ , respectively. This shared response to the value of  $\tau_D/\tau_R$  highlights the connection between the onset of nematic ordering, the non-monotonic progression of dispersity, and the block length distribution, and indicates the importance of the balance of the reaction and diffusion timescales to each of these features.

The trends observed with respect to the reaction barrier and the diffusion rate can be readily explained if we consider that a small nucleus of nascent chains grows into a crystallite by reacting with other monomers and oligomers that match the growing nuclei, both in terms of its monomer identities and in terms of its overall length. If the reaction is slowed down, relative to the relaxation time of oligomers tumbling about in the system, then the system has more time to settle into energetically favorable pre-reaction aggregates before the formation of additional polymer bonds, leading to an impressive alignment in terms of the overall oligomer length and the block lengths within each crystallite. This relative slowing of the reaction time, as compared to the system's relaxation time, can be achieved by increasing the reaction barrier or by increasing the diffusion rate, and the block length distributions shown in Fig. 5.1c&d and Fig. 5.2a clearly show these trends. As the reaction rate slows with respect to the rate of reactant diffusion, the system has additional time to relax towards lower energy configurations as chains extend. As a result, phase separation, aggregation, and nematic alignment happen earlier with respect to the reaction extent (Fig. 5.2b). This earlier aggregation and ordering leads to an earlier emergence of bonding involving longer than average chains, and a greater reduction in long chains bonding with other long chains late in the reaction (see Fig. C.7 in Appendix C for details). Such variations in bonding behaviors at different reaction extents lead to the non-monotonic dispersity behavior (Fig. 5.2c).

Together, the results in Fig. 5.1&5.2 demonstrate a sequence biasing behavior which arises from phase changes driven by the reaction itself, producing a characteristic block length which may be directly influenced through changes to the diffusivity and reactivity of the reacting species. Subsequent sections will explore the precise relationship between the chain length dependent bonding behaviors and the development of nematic and sequence

ordering in the system.

## 5.4 Characteristic block length arises from emergent changes in bonding behaviors

We now investigate the development of this sequence biasing in time, exploring how the characteristic block length arises as a consequence of a nematic ordering transition that changes the bonding behavior of oligomers within aggregates that are nucleated early in the reaction. We begin by examining, in Fig. 5.3, the time evolution of both structural snapshots and block length distributions for the reaction of stiff and flexible A,B-copolymers under our standard simulation conditions (shown in gray dashed lines in Fig. 5.1a-d and the dashed line in Fig. 5.2a).

**Early stage** ( $p \leq 0.4$ ). The simulation snapshots in Fig. 5.3a clearly demonstrate the gradual onset of aggregation, coupled with a microphase separation into **A**-rich and **B**-rich domains. The concurrent impact of these emergent aggregation processes on sequence can be seen in the time progression of the block length distribution, shown in Fig. 5.3b. At early times, the distribution of block lengths closely matches the geometric distribution expected for a polymerization process governed by Markovian statistics. However, by the time aggregates appear in Fig. 5.3a at  $20 \times 10^3 \tau$  ( $p \approx 0.4$ ), a longer tail is observed in the distribution in Fig. 5.3b, indicating the presence of longer-range ordering in the oligomer sequences – ordering that persists past the nearest sequence neighbors.

Other properties of the reaction kinetics, shown in Fig. 5.3c, are also impacted by the emergent aggregation that happens around  $20 \times 10^3 \tau$  ( $p \approx 0.4$ ): both chain-length dispersity (solid green line) and degree of polymerization (black line) increase in a non-linear manner – a sign of the breakdown in standard kinetic behaviors driven by this emergent reactant heterogeneity.<sup>58,87</sup> Longer-than-average oligomers are more prone to aggregate and therefore to encounter and react with other long oligomers, which increases dispersity and speeds chain growth. Aggregation also noticeably alters the observed probability,  $p_{AA, BB}$ , of a repeat sequence pair in the ensemble of oligomer sequences (either **AA** or **BB**). The value of  $p_{AA, BB}$  (red line) remains relatively consistent prior to aggregation, increases upon aggregation, and remains consistent at the increased value for the rest of the reaction. The impact on the sequence pair probability is a consequence of the emergence of the **A,B** microphase separation, which enriches the concentration of the like-monomer species within the local environment and leads to an increased likelihood of encounter and reaction be-

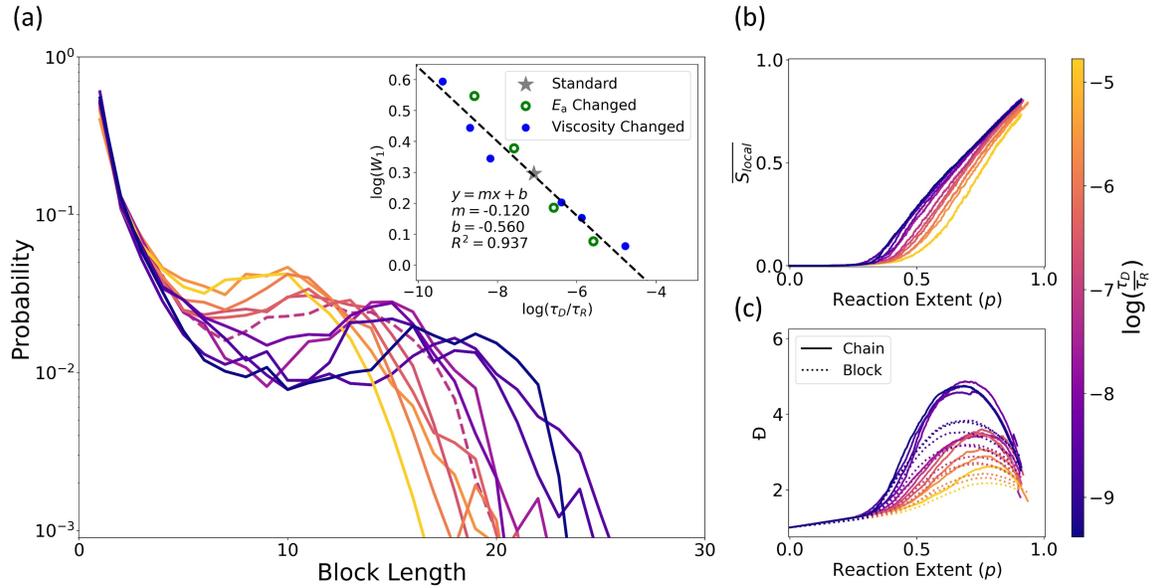
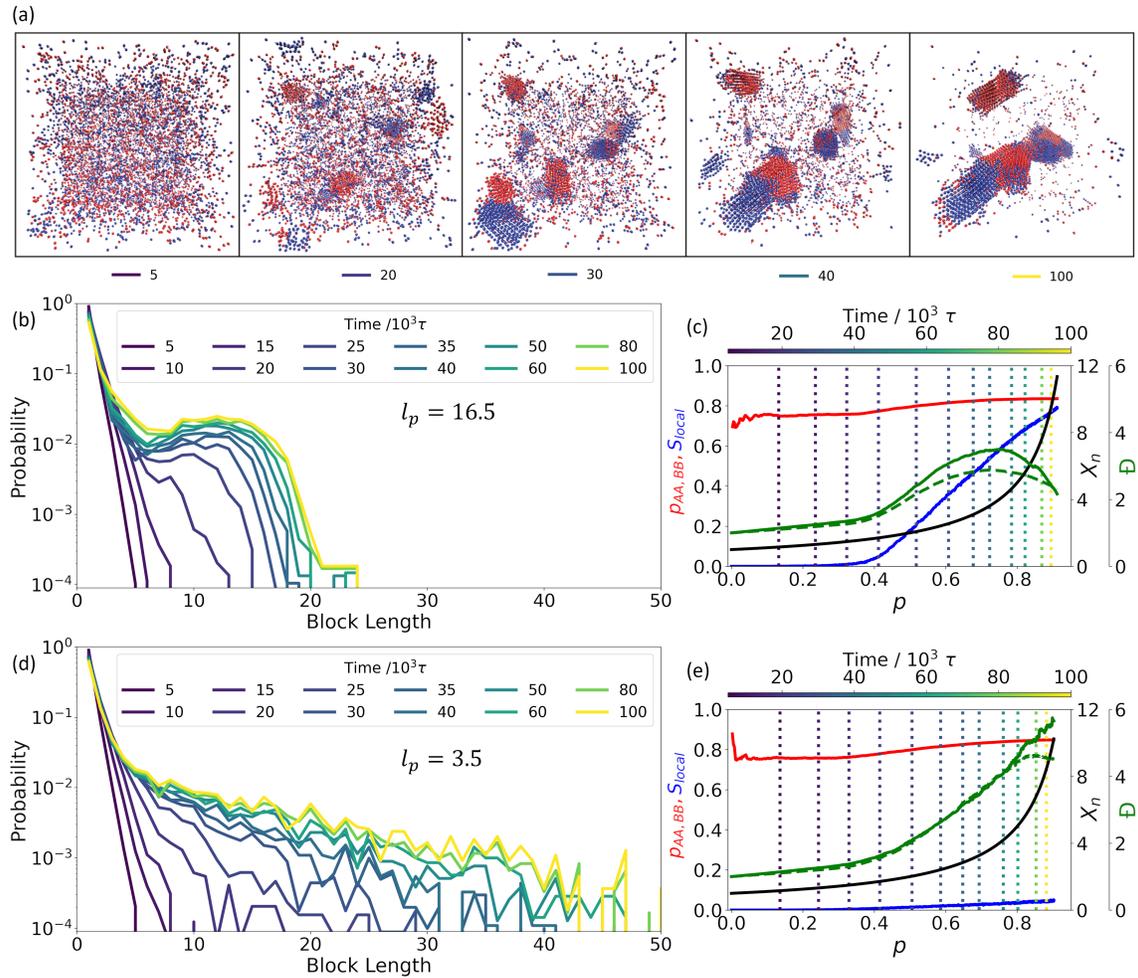


Figure 5.2: **Block length distribution shift and characteristic timescales.** (a) Block length distributions of all-A or all-B blocks observed at a reaction extent of  $p = 0.9$ , colored according to the value of  $\log(\tau_D/\tau_R)$ , the logarithm of the ratio of the characteristic diffusive and reactive timescales for the simulation. Inset: A log-log plot of the Wasserstein distance,  $W_1$ , vs. the ratio  $\tau_D/\tau_R$ . The Wasserstein distance was calculated between the observed distribution of block lengths and the distribution expected from Markov statistics (see SI for details). The result from standard simulation conditions, *i.e.*, those used in Ref. [59] and the gray dashed lines in Fig. 5.1 above, is shown in gray with a star marker. Simulations in which viscosity was varied are shown in blue, and simulations in which activation energy was varied are shown in green. A linear regression was performed, and the resulting regression equation and fit-line are shown. (b) Local nematic ordering parameter,  $\overline{S}_{local}$ , as a function of reaction extent. (c) Chain length (solid line) and block length (dotted line) dispersity,  $\mathcal{D}$ , as a function of reaction extent. All data was obtained from three independent simulation trials for each parameter set.



**Figure 5.3: System structures and block length distributions during copolymerization of semi-flexible polymers.** Results from reactive simulation of two different persistence lengths,  $l_p$ , for  $\varepsilon_{AA, BB} = k_B T$  and  $\varepsilon_{AB} = 0$ . In (a-c)  $l_p = 16.5$  monomers and in (d-e)  $l_p = 3.5$  monomers. (a) Characteristic snapshots of structures forming within a single simulation are shown at five different points in the reaction. (b,d) The block length distributions for (b) stiff chains and (d) flexible chains are plotted at twelve different time points in the reaction. (c,e) The change of sequence neighbor probability  $p_{AA, BB}$  (red line), order parameter  $\bar{S}_{local}$  (blue line), degree of polymerization  $X_n$  (black line), and chain (solid green line) and block length (dashed green line) dispersity  $\bar{D}$  are shown as functions of reaction extent,  $p$ , for (c) stiff chains and (e) flexible chains. The chosen times presented in (a), (b), and (d), are indicated with vertical dotted lines, with coloration corresponding to the simulation time.

tween monomers of the same type.<sup>58</sup>

**Middle stage** ( $0.4 < p \leq 0.7$ ). Up to this point in the reaction, both stiff chains (Fig. 5.3a-c) and flexible chains (Fig. 5.3d-e) behave in largely the same fashion, with one notable exception. For stiff chains, orientational alignment of oligomers, characterized by the order parameter  $\overline{S_{\text{local}}}$ , begins to increase with the onset of aggregation and increases continuously throughout the remainder of the reaction (Fig. 5.3c), indicating a nematic transition that is absent among the flexible chains (Fig. 5.3e). Early stiff chain aggregates are predominantly isotropic at  $20 \times 10^3 \tau$  (Fig. 5.3a), as can be seen in the low value of  $\overline{S_{\text{local}}}$  at that time (Fig. 5.3c,  $p \approx 0.4$ ). By  $30 \times 10^3 \tau$  ( $p \approx 0.6$ ), however, the chains within each aggregate show signs of significant ordering (Fig. 5.3a), with  $\overline{S_{\text{local}}}$  continuing to increase as chains align and the crystallites grow.

Concurrent with this nematic ordering, between  $20 \times 10^3 \tau$  and  $30 \times 10^3 \tau$ , a peak forms in the block length distribution of the stiff chains (Fig. 5.3b), indicating the emergence of a characteristic block length, which remains absent from the equivalent distribution of flexible chains (Fig. 5.3d). This distributional peak shifts to the right and becomes more distinct as the reaction proceeds, and by  $40 \times 10^3 \tau$  ( $p \approx 0.7$ ) the characteristic block length observed at the end of the reaction is well established.

**Late stage** ( $0.7 < p \leq 0.9$ ). Additional bonding past this stage generally forms more sequences of that characteristic length rather than increasing it, so that the distribution curves begin to coalesce. Importantly, the decrease in dispersity shown in Fig. 5.3c, also occurs at this point, arising as a consequence of the nematic alignment once sufficient ordering occurs in the system ( $\overline{S_{\text{local}}} \gtrsim 0.5$ ). This decrease indicates that chains that are longer than average are preferentially reacting with chains that are shorter than average, reducing the overall dispersity. This two-stage, length dependent bonding behavior is unique to the aggregates formed from stiff chains. Flexible chains demonstrate no such decrease in dispersity, and the preference for long oligomers to bond other long oligomers, which spurred the initial non-linear increase in dispersity, continues throughout the reaction (Fig. 5.3e).

**Chain length dependent bonding.** To further quantify the length dependent bonding behavior we observe, we present in Fig. 5.4 histograms of the likelihood of bonding between chains of different lengths over the early ( $p \leq 0.4$ , top row), middle ( $0.4 < p \leq 0.7$ , middle row), and late ( $0.7 < p \leq 0.9$ , bottom row) stages of the reaction. Fig. 5.4a shows the results from trials of a simple, idealized Monte Carlo calculation designed to reproduce the sequence statistics expected under conditions where Flory's equal reactivity principle<sup>73</sup> holds – namely, where all extant chains are equally likely to form a bond (see Chapter 2 for details). While such a calculation fundamentally cannot account for spatial heterogeneities

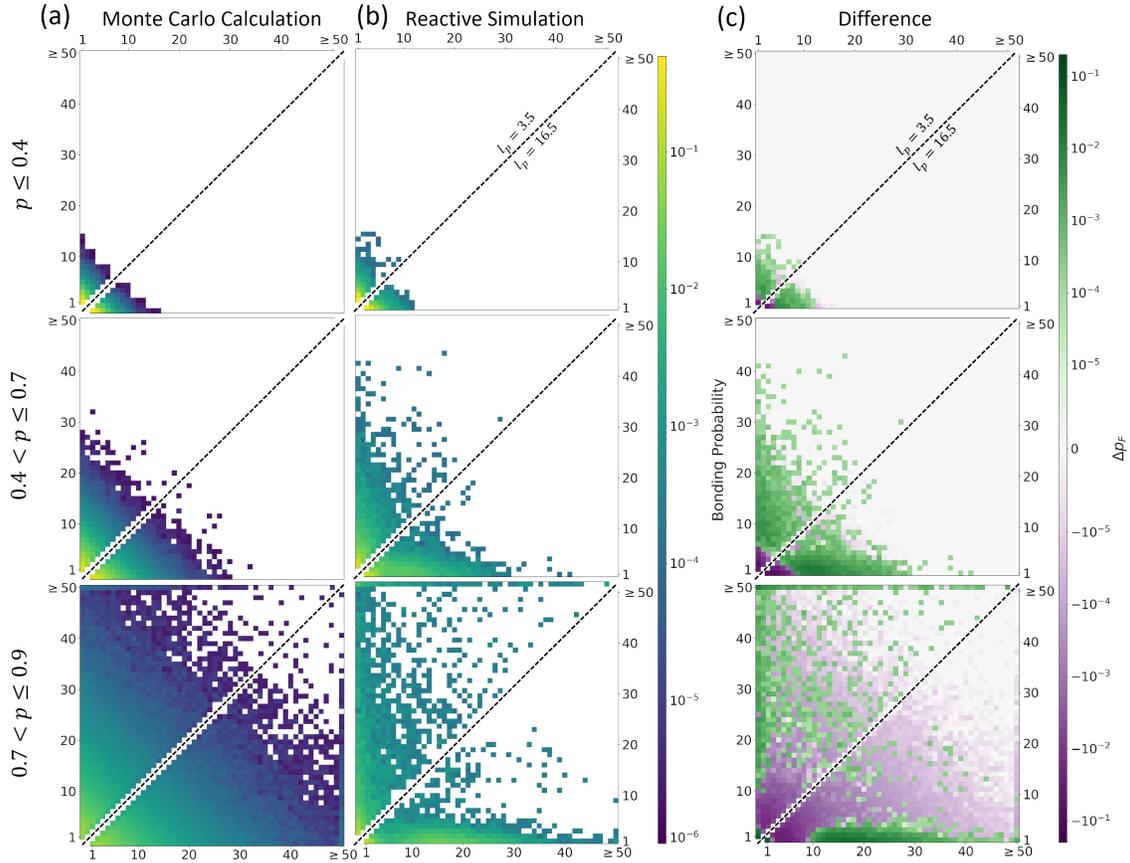
of reactants or the stiffness and orientations of the nascent chain, it is nevertheless an effective implementation of the assumptions of the equal reactivity principle, which similarly does not capture these details. Fig. 5.4b shows the results for our model system under both flexible (upper-left diagonal) and stiff (lower-right diagonal) cases. Fig. 5.4c shows the difference,  $\Delta p_F \equiv p_{\text{observed}} - p_F$ , in bonding probability between our observations,  $p_{\text{observed}}$ , and the ideal Flory behavior,  $p_F$ .

Under equal reactivity (Fig. 5.4a), early stage bonding is dominated by monomer-monomer reactions, and only late in the reaction do longer chains begin to emerge. Even late in the reaction, the majority of bonding involves short chains of length  $\leq 3$ , and only very rarely do two longer oligomers react. It should be noted here that there is no difference in bonding probabilities arising from differences in chain stiffness, as by definition equal reactivity, which is taken as the basis of these Monte Carlo calculation, does not permit such properties of the larger chain to affect the bonding behavior. The chain persistence length therefore does not enter into this calculation, and the equivalent behavior of each side of the diagonal line in these plots is axiomatic.

The ideal Flory bonding behavior in Fig. 5.4a contrasts sharply with the results from our reactive Langevin simulations in Fig. 5.4b, and the difference is quantified in Fig. 5.4c. In the dynamical simulations, early stages of the reaction show increased formation of longer chains and reduced monomer-monomer bonding as compared to the ideal behavior, a consequence of the aggregation spurred by chain lengthening. The propensity of longer chains to aggregate causes them to more readily encounter and react with longer chains and produces the length dependent increase to bonding, deviating from Flory's ideal behavior. For the early stage of the reaction ( $p \leq 0.4$ ), these behaviors are largely independent of chain stiffness, as chain lengths well below the persistence length are insufficient to drive the nematic ordering transition which occurs later in the reaction for stiff chains. As the reaction proceeds further ( $0.4 < p \leq 0.7$ ), the chain length dependence of bonding probabilities begins to differ between chains of different persistence length, as chains with sufficient stiffness begin to nematically align. Over this reaction extent range the difference is small, with flexible chains showing only a slight increase in the likelihood of bonding between longer chains. It is only in the late stage of the reaction ( $0.7 < p \leq 0.9$ ), when substantial nematic ordering has occurred, that greater differences in the bonding behaviors between flexible and stiff chains are observed. For flexible chains at this stage, there is an increased likelihood of reaction as compared to the ideal Flory behavior for a wide range of oligomer chain length combinations. Flexible chains also have a particularly noticeable enrichment of bonding events that involve very long oligomers (length  $\geq 50$  monomers), as seen in the

band of green along the top of the bottom plot in Fig. 5.4c. By contrast, stiff chains show a marked increase of short chains (length  $\leq 3$  monomers) bonding with longer chains over this reaction extent, as visualized by the band of green along the  $x$ -axis. Both flexible and stiff chains also show a marked reduction in the bonding between shorter oligomers, visible in the purple region in the lower left. To further quantify these differences and how they develop as a function of reaction extent, in Fig. C.2 in Appendix B we report the relative proportion of bonding involving at least one chain of length  $\geq 3$  monomers and of bonding between both chains of length  $\geq 3$  monomers, for both flexible and stiff chains.

Taken together, the results in Fig. 5.4 show that the transition to nematic ordering among the stiff chains, seen in the increase of the  $\overline{S}_{\text{local}}$  metric near  $p = 0.4$  in Fig. 5.3c, occurs concomitantly with a change in chain-length dependent bonding behavior: a transition from longer chains preferentially reacting with longer chains, to longer chains preferentially reacting with shorter chains. It is this shift in bonding behavior which results in the non-monotonic dispersity we see in Fig. 5.3c. When this nematic ordering and the associated reduction in dispersity occurs alongside the polymerization driven phase separation into **A**-rich and **B**-rich domains, peaked block length distributions emerge with a characteristic block length (Fig. 5.3). Once chains have aligned, it becomes energetically favorable for oligomers shorter than the length of the growing crystallite to orient themselves such that their combined length spans the crystalline domain. Once in that position, a reaction between these aligned oligomers becomes more likely. We term this class of kinetics self-templating growth, and examine it in the context of the bonding behaviors of nucleated oligomers in the following section.



**Figure 5.4: Chain-length dependence of bond formation.** These plots show the probability that chains of different lengths will form a bond during different periods of the reaction. Each row corresponds to bond events sampled during a different range of the simulated reaction extent:  $p \leq 0.4$ ;  $0.4 < p \leq 0.7$ ; and  $0.7 < p \leq 0.9$ . Within each plot in (b) and (c), the upper-left half shows the values for flexible chain systems ( $l_p = 3.5$ ) and the lower-right half shows the values for stiff chains ( $l_p = 16.5$ ). Probabilities in Column (a),  $p_F$ , were determined from 250 simple Monte Carlo calculations of the ideal Flory behavior, in which no spatial or stiffness information is taken into account, and thus all chains have an equal probability of reacting at each step in the reaction and the two halves are therefore mirror images. Probabilities in Column (b),  $p_{\text{observed}}$ , were determined from actual bonding events observed in our reactive Langevin dynamics simulations from three independent trials for both flexible and stiff chains. In Column (c), the difference,  $\Delta p_F$ , is shown between the observed probability in (b) and the probability which occurs under ideal, Flory behavior in (a). All probabilities are normalized by the the number of bonding events within the stated reaction extent range.

## 5.5 Oligomer aggregate structures and bonding behaviors depend on chain stiffness

**Quantitative identification of oligomer aggregates.** In order to examine this self-templating growth further, we now consider the initial aggregation, alignment, and growth of individual crystallites in the system. To this end, we first develop a consistent method of identification of aggregated regions in which the transition to nematic order and self-templated growth may occur. We therefore make use of a modification of the HDBSCAN clustering algorithm,<sup>122</sup> which identifies spatially distinct oligomer clusters based on a local density criteria (see Sec. C.4 in Appendix C for details). Clusters are initially identified as regions with a minimum of twelve monomers within a distance of  $2.5 \sigma$ . In addition, when one or more monomers of a chain are identified as belonging to a cluster, the entire chain is then added to that cluster. These conditions for cluster identification were chosen so that both small nucleated regions forming early in the reaction, as well as larger, spatially separated aggregates occurring later in the reaction, can be identified via consistent clustering criteria.

**Aggregate growth and composition.** The results of the density based clustering analysis are presented in Fig. 5.5. In Fig. 5.5a, the growth of clusters identified within the simulation is visualized at different points throughout the reaction for both stiff and flexible chains. Irrespective of chain stiffness, no clusters are identified early in the simulation. Only at times  $\geq 10 \times 10^3 \tau$ , where sufficient polymerization has occurred to drive assembly, do the first small clusters form. From  $15 \times 10^3 \tau$  to  $20 \times 10^3 \tau$  ( $p \approx 0.3$  to  $p \approx 0.4$ ) the number of identified clusters increases rapidly, as the reaction proceeds further and multiple oligomer clusters nucleate. Further aggregation proceeds through two pathways, with aggregates growing both from the addition of unaggregated chains from the surrounding, dilute phase, as well as from the merging of separately nucleated clusters. This process of aggregate nucleation and subsequent growth and merger is quantified in Fig. 5.5b, in which the number of distinct aggregates increases sharply during the initial nucleation phase before steadily decreasing throughout the remainder of the reaction as aggregates merge. Meanwhile, the total proportion of all monomers belonging to an aggregate increases continuously, as monomers and chains not belonging to aggregates bond and join with existing aggregates. Notably, this behavior appears independent of persistence length, as both flexible and stiff chains show nearly identical aggregate nucleation, growth, and cluster merging dynamics.

In Fig. 5.5c, we consider the fractional comonomer composition of the clusters through-

out the reaction. Due to the influence of the non-bonded attractions between like-monomer species, almost all aggregates are initially dominated by one comonomer species (Fig. 5.5c). Of course, due to the relatively small size of clusters in the early stages of formation, some outliers in cluster composition occur early on where the comparatively rare the inclusion of few monomers of the non-dominant species in a cluster will have a disproportionate impact on the cluster composition as a whole. As the reaction continues and mergers between **A**-rich and **B**-rich aggregates become more frequent, the composition fraction starts to even out.

As seen in prior work<sup>58,86</sup> and discussed in Section 5.2, the combination of local density increase and composition shift leads to a shift towards longer sequence block lengths than would be present in a copolymerization which remains homogeneous throughout the reaction. This phenomenon is borne out within the individual aggregates, which possess peak block lengths (Fig. 5.5d) which correspond to the largest block lengths observed in the system as a whole (Fig. 5.3b), much longer than that of the unaggregated sequences in the system. Block lengths of unaggregated chains remain  $\leq 3$  monomers in length throughout the reaction (see Fig. C.5 in Appendix C), confirming that the increased block lengths observed in the system are isolated to the aggregated regions we identify. Significantly, although the initial nucleation, growth rate, and composition of aggregates are largely independent of chain stiffness (Fig. 5.5b&c) the growth of contiguous sequence blocks within the aggregates differs between chains of different persistence lengths (Fig. 5.5d).

**Chain stiffness influences bonding within and between aggregates.** In order to better understand this apparent contrast between similar aggregate growth pathways and dissimilar sequence development, we now consider the source of additions to growing aggregates. To this end, we consider three distinct populations of bond formation events for a given aggregate: (1) “unclustered” bonds, where a bond forms between a chain in the aggregate and a chain that is not part of any aggregate; (2) “self” bonds, where a bond forms between two chains that have been part of the same aggregate for at least  $5 \times 10^3 \tau$ ; and (3) “merge” bonds, where a bond forms between two chains that are part of separate, distinct aggregates.

Fig. 5.6 shows the results of this bonding population analysis, performed on the same flexible and stiff chain systems shown in Fig. 5.5. The most readily observable difference in bonding behavior between flexible and stiff chain aggregates lies in the number of unclustered bonding events (Fig. 5.6a), as the stiff chain aggregates form more bonds with the unclustered population of chains than the flexible chain aggregates do. This change in bonding behavior is critical: unclustered chains have shorter chain and block lengths

than those in aggregated regions (Fig. C.4&C.5 in Appendix C), since they have not been influenced by the emerging phase-separation and assembly effects. The unclustered chains with shorter than average chain and block lengths bond with aggregated chains with longer than average chain and block lengths, narrowing the respective distributions towards the mean. Increased incorporation of these unclustered chains in the stiff chain case therefore serves to reduce the dispersity of both chain and block lengths and slow the distribution broadening caused by aggregation. For bonding occurring between chains within the same cluster (Fig. 5.6b), the opposite relationship with persistence length is observed. Here, stiff chain aggregates demonstrate reduced self bonding behavior as compared to flexible chain aggregates after the transition to nematic ordering has begun ( $p > 0.4$ ). Thus, for flexible chains, bond formation continues to occur preferentially between chains within the aggregated population. Due to the longer chain and block lengths of that population, this propensity towards self bonding causes chain length and block length dispersities to continue to increase. Notably, this difference in bonding behaviors of aggregates of different chain stiffness is not seen in bonding events associated with the merger of separate clusters (Fig. 5.6c). Bonds from merging clusters make up a smaller portion of the total bonding events observed, and the quantities of these events, and the timing of their occurrence is relatively consistent between both stiff and flexible chains.

**Aggregate structure alters accessibility of reactive chain ends.** In considering the impact of chain stiffness on bonding within and between aggregates, we hypothesized that both of the above observations of altered aggregate bonding behavior are related to the relative positioning of chain ends within aggregate structures. In flexible chains, and in stiff chains prior to nematic ordering, chain ends may be isolated in the interior of an aggregate, where they are less readily accessible to chains which are not a part of the same aggregate. In contrast, the ordering transition that stiff chain aggregates undergo positions chain ends in the exterior of the aggregated structure. This positioning not only increases the accessibility of chain ends to the surrounding local environment of the aggregate, but also limits the contact between the inflexible chains of the same aggregate, simultaneously promoting bonding from outside the aggregate and limiting bonding from within it. The timing of this ordering transition for stiff chains is therefore critical, as early aggregation first promotes distribution broadening by increasing bonding between the longer chains within an aggregate, before then promoting bonding with the shorter chains of the surrounding environment after nematic alignment occurs.

In Fig. 5.7, we quantify the accessibility of all monomers and chain end monomers, within aggregates of both stiff and flexible chains. When considering all aggregated monomers

in Fig. 5.7a, we see that the fraction of monomers belonging to either the interior or exterior of an aggregate increases with reaction extent. The largely spherical aggregates produced by flexible chains have more of their aggregated monomers in the interior, as compared to the more extended aggregates of stiff chains, which have more of their aggregated monomers in the exterior. The average number of local (within  $2.5\sigma$ ) neighbors,  $n^{\text{local}}$ , in the system is plotted in Fig. 5.7b for all monomers. As expected, the local density increases monotonically with reaction extent due to the polymerization driven aggregation. Despite their more extended character, however, the stiff chains demonstrate a higher average density after nematically ordering – a consequence of the closely packed structures such alignment produces.

When considering the fraction of chain end monomers in aggregate exteriors and interiors in Fig. 5.7c, we find that chain ends are found exclusively in the exteriors of stiff chain aggregates. In contrast, an appreciable number of chain ends remain in the interior region of flexible aggregates, and this number grows steadily as the reaction proceeds and the interior volume increases. These differences can also be seen in the  $n^{\text{local}}$  of the chain ends in Fig. 5.7d. Although the curves initially look the same as those for all monomers (Fig. 5.7b), a cross-over occurs at  $p \approx 0.4$ , when the stiff chains begin to nematically order, after which the flexible chain ends show greater neighbor density than their stiff chain counterparts. Only very late in the reaction,  $p \gtrsim 0.8$ , does the local density near stiff chain ends again surpass those of flexible chains.

Overall, the nematic ordering in stiff chains ensures that chain ends, where new bond formation can occur, are restricted to the less dense exterior region of aggregates (Fig. 5.7c&d) where they are more readily accessible for bond formation with shorter, unaggregated chains (Fig. 5.6, Fig. C.4 in Appendix C).

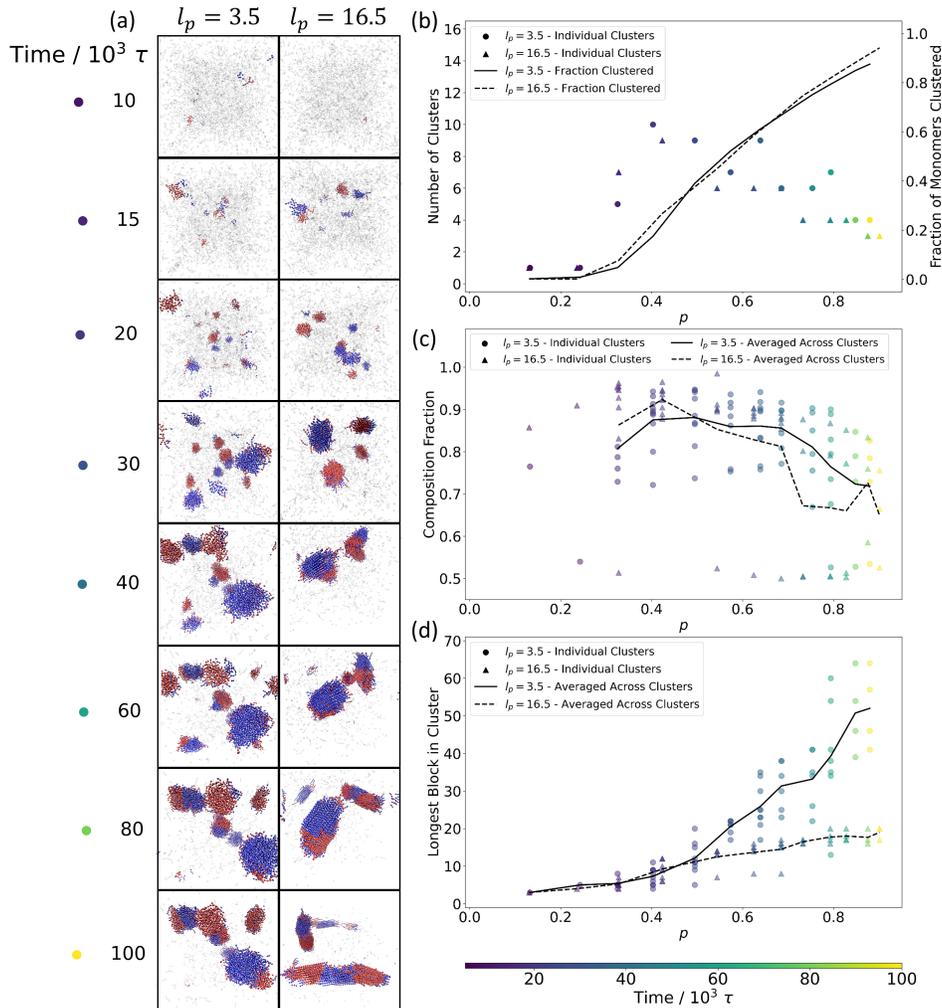


Figure 5.5: **Identification of spatially clustered aggregates and their growth.** (a) Aggregated structures identified by clustering analysis for both flexible ( $l_p = 3.5$ ) and stiff ( $l_p = 16.5$ ) chains. Monomers belonging to identified clusters are shown in color, with species A in red and species B in blue, while all other monomers are shown in transparent gray. (b) The number of separate clusters (points, left axis) and the fraction of all monomers belonging to a cluster (lines, right axis), are plotted vs. the reaction extent,  $p$ . (c) The composition fraction of the dominant monomer species within a cluster is plotted as a function of the reaction extent,  $p$ . Each point corresponds to a single identified cluster at that reaction extent. Lines show the average value over all identified clusters. (d) The longest block length among all the sequences belonging to a single cluster is plotted here as a function of reaction extent,  $p$ . Each point corresponds to a single identified cluster at that reaction extent. Lines show the average longest block value over all identified clusters. In (b-d) marker coloration indicates simulation time, while marker style indicates chain stiffness.

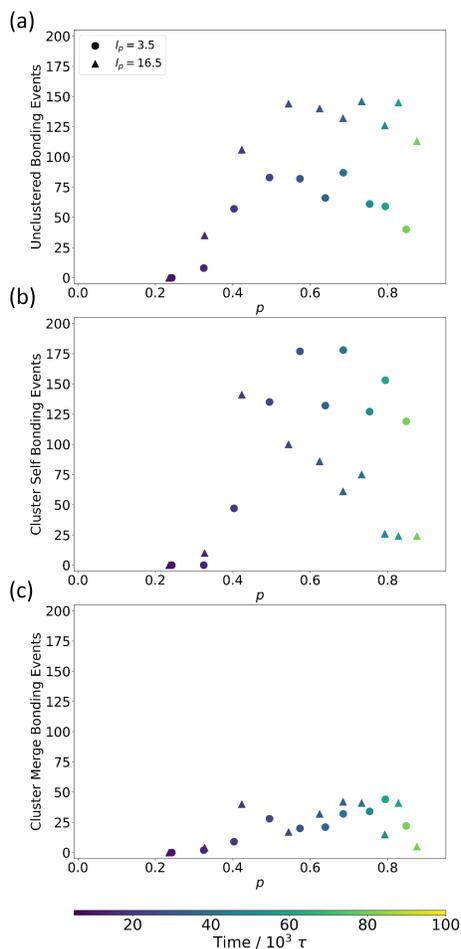


Figure 5.6: **Oligomer bonding in spatially clustered aggregates.** The number of oligomer bonding events occurring within aggregates identified by clustering analysis as a function of reaction extent for both flexible ( $l_p = 3.5$ ) and stiff ( $l_p = 16.5$ ) chains. Bonding events were split into three populations: (a) bonding with unclustered reactants, (b) bonding within a single cluster, and (c) bonding between separate clusters. Each data point corresponds to the total number of bonding events of each type among all clusters in the system at the indicated reaction extent. Time is indicated by marker coloration and chain stiffness is indicated by marker style.

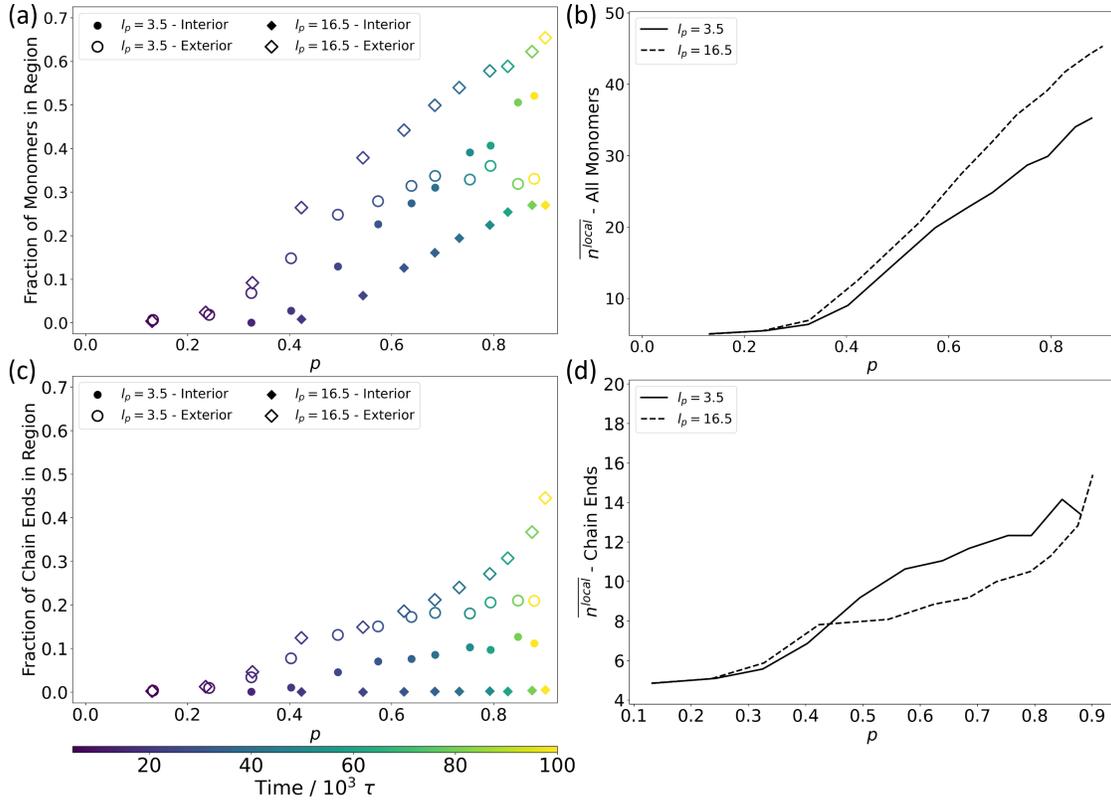


Figure 5.7: **Distribution of chain ends within aggregates.** (a) The fractions of all monomers in the system that belong to an interior or exterior region of an aggregate are plotted as a function of the reaction extent,  $p$ , for both flexible ( $l_p = 3.5$ ) and stiff ( $l_p = 16.5$ ) chains. Aggregate regions were identified by a local density criteria using  $n^{local}$  to quantify the number of neighboring monomers within  $2.5\sigma$ . Monomers with  $n^{local} \geq 36$  were identified as belonging to the interior of aggregates (filled markers), while monomers with  $12 \leq n^{local} < 36$  were identified as belonging to the exterior of aggregates (hollow markers). Marker style indicates chain stiffness, and coloration indicates simulation time. (b)  $\overline{n^{local}}$ , the average number of local neighbors for all monomers in the system, is plotted vs. reaction extent for flexible and stiff chains. (c) The fractions of all chain ends in the system that are located within an aggregate's interior and exterior regions are plotted as a function of reaction extent,  $p$ , for both flexible and stiff chains. (d) The average number of neighbors in the local environment,  $\overline{n^{local}}$ , is plotted here for all chain ends in the system for both flexible and stiff chains.

## 5.6 Summary and conclusions

In this work we have investigated a previously reported<sup>59</sup> templating effect in step-growth copolymerizations of monomer species with sufficient chain stiffness and non-bonded interactions between like species, an effect which results in a characteristic block length of sequence repeats. Here, we have examined the response of this characteristic block length to changes in reaction condition, demonstrating its sensitivity to changes to the relative timescales of diffusion and reaction, such as those introduced by changes to solution viscosity or reaction activation energy. We have shown, in Fig. 5.2, that the characteristic block length shifts in response to changes in these timescales in a quantifiable fashion, suggesting a pathway for biasing copolymer sequences towards a particular block length through understanding this effect.

Exploring the development of the templating effect in time shows that the emergence of the characteristic length coincides with a phase separation and nematic ordering in the system unique to stiff chains (Fig. 5.3), and that this drives changes in length dependent bonding behavior associated with these transitions (Fig. 5.4). It is the simultaneous impact of these effects which together produce the unique characteristic block length behavior we observe. These results further contextualize and reinforce the impact of characteristic system timescales described in Fig. 5.2. Because the characteristic block length develops as a consequence of both phase separation and nematic ordering, changes to the relative timescales of diffusion and reaction which govern the rates of these processes necessarily alter the point at which self-templating growth begins, the extent to which it influences sequences, and thereby the characteristic block length that emerges.

To further understand how the combination of nematic ordering and phase separation introduces length dependent bonding behaviors in stiff chained oligomers, we have contrasted the growth pathways of aggregated oligomer structures of flexible and stiff chains (Fig. 5.5). We have highlighted the differences in bonding behaviors of these aggregates in terms of the reactant populations which spur their growth (Fig. 5.6, and motivated this in terms of the relative accessibility of unreacted chain ends to these populations (Fig. 5.7). Together, these results explain the unique chain stiffness and chain length dependence of bonding described in the previous sections (Fig. 5.4), and support our description of the impact of characteristic system timescales on the phenomenon of a characteristic block lengths that we observe (Fig. 5.2). The structure of aggregates and their growth in time is what is fundamentally altered by nematic ordering, and it is this change, in concert with the ongoing phase separation in the system, that leads to the unique bonding behaviors we ob-

serve. Changes to characteristic timescales in the system alter when these aggregates form and when they order, relative to the rate at which polymerization proceeds and sequences and chain lengths develop, promoting or truncating the impacts of aggregation and ordering on these features.

Our results have demonstrated how the templating effect arises from the simultaneous impacts of polymerization-driven phase separation and nematic transition which alter the bonding behavior of nascent chains in a length-dependent fashion. These length-dependent bonding behaviors depend on the formation, growth, and structural ordering of oligomer aggregates, which changes the accessibility of the aggregated oligomers to other populations of reactants. Altering the characteristic reactive and diffusive timescales in the system changes the timing of formation of aggregates and their ordering with respect to the progress of the reaction, changing the extent to which the chain-length and chain-stiffness dependent bonding behaviors may influence sequence development, and thereby shifting the resultant characteristic sequence development that emerges.

# Chapter 6

## Conclusions

### 6.1 Summary of results

We have utilized a particle based, bead-spring model, sampled via a reactive Langevin dynamics approach, to simulate a two component, step-growth copolymerization under a variety of reaction conditions. This method has allowed us to directly observe the development of oligomer sequences, and correlate them to reaction kinetics and emergent reactant heterogeneity, in a fashion not possible with current experimental characterization techniques for polymerization reactions. With this approach, we have investigated interaction regimes which lead to emergent self-assembly and microphase separation of reactants, and studied the associated impacts on reaction kinetics, oligomer sequence, and end product dispersity and morphology.

In Chapter 3, we explored the combined effects on copolymerization kinetics of differences in both activation energies and non-bonded attractions between monomers and examine the sequences produced within this same step-growth copolymerization model, connecting our results to classical kinetic theories of copolymerization. We have demonstrated that our model successfully reproduces kinetic and sequence behaviors predicted by current theories and that we can quantitatively predict how altering activation energies and non-bonded attractions influences the kinetics and sequences. Critically, we also demonstrated a regime in which non-bonded attractions are sufficient for polymerization induced assembly, leading to non-standard kinetic behaviors and long-range sequence biasing, with the extent of each clearly shifting as the reaction proceeds. We showcased how these impacts of reactant assembly are not captured by existing theories and that they depend on

both non-bonded attraction strength and reaction activation energies, exhibiting that the influence of each of these energies may work in concert or in opposition to one another to bias the sequences formed.

In Chapter 4 we explored the emergent assembly phenomenon further, examining how effective attractive interactions, mediated by a solvent selective for one of the reacting species, impact the development of sequence and the supramolecular assembly in a simple A-B copolymerization. We found that as the effective attractions between monomers increase, an emergent self-organization of the reactants causes a shift in reaction kinetics and sequence development. When the solvent-mediated interactions are selective enough, the simple mixture of A and B monomers oligomerize and self-assemble into structures characteristic of amphiphilic copolymers. The composition and morphology of these structures and the sequences of their chains are sensitive to the relative balance of affinities between the comonomer species.

In Chapters 3&4, and in previous work,<sup>58,86,87</sup> we have observed that polymer chains with sufficient stiffness and intermolecular attractions can undergo an emergent, polymerization-driven nematic alignment of nascent oligomers during a step-growth polymerization process. Both the extent of alignment and the point in the reaction at which alignment occurs impacts the kinetics and the sequence development of the growing polymer. Of particular interest is the emergence of a characteristic block length in the ensemble of sequences, resulting in unusually peaked block length distributions. In Chapter 5 we explored the emergence of this characteristic block length and investigate how changes in activation energy, solution viscosity, and monomer density influence the sequence and block length distributions of stiff copolymers undergoing step-growth polymerization. We found that emergent aggregation and nematic ordering restricts the availability of longer chains to form bonds, altering the propensity of chains to react in a length dependent fashion which changes as the reaction progresses, promoting the formation of chains and blocks of a characteristic length. Further, we demonstrated that the characteristic length scale which emerges is sensitive to the relative timescales of reaction kinetics and reactant diffusion, shifting in response to changes in the activation energy of the reaction and the viscosity of the solvent.

## 6.2 Impact and applications

These results presented in this work extend previous observations of kinetic alteration and sequence biasing effects in step-growth polymerizations<sup>58,59</sup> to cases where effective in-

termonomer attractions or reaction activation energies are asymmetric, such as would arise when solvent affinities differ between reactants or when the chemical identity of a monomer influences its reactivity. These conditions relate to what is generally expected for real-world polymerizations, where such differences in reactant affinities and monomer reactivities exist.<sup>113–115,123–125</sup> Localized concentration enrichment arising from solvent selectivity, of the kind we observe here, has been proposed as a mechanism for the “bootstrap” effects discussed previously<sup>57,60,61</sup> and for the reaction rate increases in PISA formulations.<sup>66,69,70</sup>

Our analysis of oligomer sequences shows that, for the more attractive (i.e., solvophobic) **A** monomers, the contiguous **A**-block length distributions show a comparable shift to that of the overall chain length distributions, demonstrating that conditions which foster concentration inhomogeneities also impact the resulting oligomer sequences. For conditions in which monomer affinities differ sufficiently, even the relative incorporation of **A** and **B** species into the aggregated phase is altered. This alteration changes the development of sequences, promoting the formation of longer block lengths of the more aggregated monomer species.

The differential incorporation of monomer species into the aggregate phase highlights another important consequence of solvent selectivity. In systems with equivalent attractions between all monomer species, both species have the same propensity for aggregate formation, however asymmetry in the attractions affecting the different species leads to aggregates with distinct domains of altered composition. Aggregates of monomers with sufficiently asymmetric attractions show compositions comparable to micellar or core-shell structures common in supramolecular self-assembly and emulsion polymerizations.<sup>66,126,127</sup> This behavior demonstrates how the balance of attractive interactions between reactants can impact not only the sequence of oligomers and the morphology of aggregates, but also the inclusion and distribution of species throughout the aggregate structures. The combined influence of sequence and solvent quality on copolymer aggregates has also been shown in computational studies of block, statistical, and gradient copolymers.<sup>128–130</sup> Our results reinforce this connection, further demonstrating how solvent interactions may also influence the early development of sequence. In addition, we find that these effects are altered by the geometric constraints associated with stiffer chains – as expected given prior studies on the solvent-dependent phase behavior of stiff and liquid crystalline polymers.<sup>120,131–133</sup>

The balance of solvophobic and solvophilic interactions is a critical component in PISA, which provides a notable example of both the importance and utility of the dynamic interplay between the relative affinities of reactants and solvent, as well as a multi-step assembly pathway which is initiated by the formation of strong anisotropic bonds. Recent extensions

of PISA processes to sequence controlled polymers suggest the opportunity for improved morphological control, highlighting the potential for additional development and the need for greater understanding in this area.<sup>134,135</sup> Further work has suggested that supramolecular morphologies accessible through PISA may be impacted by the addition of grafted side chains, which can alter morphology in a sequence-dependent manner.<sup>136,137</sup> Other polymerization induced microphase separations have been shown to promote desirable material properties as well,<sup>138,139</sup> and there is continued interest in the use of solvent interactions to tune supramolecular polymer structures.<sup>123</sup>

Alongside our own results, these observations of the importance of solvent interactions to copolymer phase behavior highlight the potential to simultaneously bias sequence and tune aggregate morphology and composition through the intentional alteration of monomer-solvent affinities. Even the strongest set of attractions explored in this work,  $\epsilon_{\text{att}} = k_{\text{B}}T$  (which is  $\sim 2.5$  kJ/mol at 300 K), are readily accessible interaction energies for a variety of supramolecular interactions,<sup>140–142</sup> making it feasible to select or modify comonomers and solvent to obtain desired interaction profiles. The persistence lengths we explore are also well within the range observed in flexible and semi-flexible chains of both biological and synthetic polymers.<sup>143–147</sup> Interestingly, according to Flory-Huggins theory,<sup>73</sup> even the milder attractions of  $\sim 0.3k_{\text{B}}T$ , which do not exhibit an emergent phase separation in our simulations, would still be expected to spur aggregation at some longer length. Tailoring the length of the copolymers at the onset of this behavior could provide another way to modify sequence and assembly.

The interplay between solvent interaction, sequence, and aggregation behavior suggests the possibility of tuning the relative solvent affinities of the reactants in order to intentionally bias the resulting copolymer sequence and assembly. The control of solvent interactions may be a feasible route to control phase behavior and material properties of copolymers. One recent study of note<sup>148</sup> makes use of precisely such phenomenon, tuning the pH in an aqueous dispersion minima to produce populations of copolymers with a distinct composition profile, and noting a strong associated impact on their thermoresponsiveness.

More generally, the assembly processes we observe in this work belong to the broad class of non-classical self-assembly pathways.<sup>149–153</sup> In many of these non-classical assembly processes, weaker isotropic interactions drive the formation of a condensed, but still disordered, phase, after which anisotropic interactions among the condensed particles give rise to further ordering, as seen for instance in models of protein crystallization and in the formation of "shish-kebab" polymer crystallites around a fiber-like core.<sup>154–157</sup> This "two-step" assembly pathway describes well what we observe in these polymerizations, once

the reaction has yielded oligomers of sufficient length to aggregate. After reaching this threshold length, the nascent oligomers condense and the interplay between the isotropic selective solvent-mediated interactions and the orientationally-dependent oligomer bonds can give rise to additional nematic and sequence ordering. However, in our system this "two-step" pathway is itself initiated by the formation of the strong anisotropic polymer bonds that act to reduce the entropic cost associated with monomer aggregation and thus shift the thermodynamics of the system towards the formation of an initially less-ordered condensed phase.<sup>73</sup>

Previous work has posited similar length matching or length homogenizing effects in aligned polymers with non-bonded attractions between neighboring chains, and we observe such a cooperative aggregation effect in our own system, biasing the formation of structures with fewer "loose ends".<sup>158</sup> Though distinct in mechanism, the "gel effect"<sup>159,160</sup> also produces autocatalytic kinetics and length dependent bonding behaviors<sup>161</sup> in radical polymerizations analogous to the behaviors we see here, through emergent and length-dependent changes to the diffusion of reactants. In the gel effect, reduced diffusion of longer chains slows the rate of the termination step, allowing the competing propagation step to occur more rapidly and increasing the polymerization rate. By contrast, in our system, autoacceleration and length dependent bonding arise instead from non-bonded interactions and the associated aggregation and demixing of chains of sufficient length. It is in the subsequent growth of the aggregated phase in which templating behavior occurs and the characteristic block length emerges.

The results presented in this work are of particular interest in the context of rod-rod conjugated copolymers,<sup>162,163</sup> in which both comonomer species exhibit backbone rigidities analogous to the stiff chain copolymers explored here.<sup>164</sup> Rod-rod conjugated copolymers have demonstrated a number of optoelectronic material properties of interest,<sup>162,165–167</sup> and crucially, these properties have been shown to be strongly influenced by the primary monomer sequence and the size of respective blocks.<sup>168–170</sup> Further, these types of "all conjugated" copolymers are frequently developed through step-growth polymerization processes,<sup>171–174</sup> the same reaction schema we model throughout this work. The collective development of characteristic sequence repeats we observe provide a potential avenue for directly biasing copolymer sequence in a wide range of chemical systems through the informed modification of the balance of reactive and diffusive timescales of the reaction. Further, the behavior of oligomeric aggregates of conjugated polymers have attracted attention as an avenue to both improve solvent processability and to tune material properties through mesoscale structure.<sup>175–178</sup> Our results provide insight into the development of oligomeric

aggregates in stiff chained copolymer systems, showcasing the interplay between aggregate structure and ordering and the resulting sequence and molecular weight distributions of growing oligomers.

Finally, though the motivations of this work aimed to explore sequence development and sequence influencing effects in non-biological polymers, there are some relevant connections to biopolymers as well. These connections arise from the particular observation of collective reactant behaviors and emergent reaction heterogeneity caused by polymerization in an initially homogeneous system leading to a breakdown in the expected Flory-Schulz chain length statistics and the development of longer than expected chains at a given reaction extent. Such an observation may be of particular interest in the context of the so-called “Flory length problem”<sup>179</sup> in the study of the chemistry to biology transition of prebiotic polymers. Previous work has demonstrated that RNA,<sup>180</sup> peptides,<sup>181,182</sup> or other model prebiotic polymers,<sup>183–185</sup> can form autocatalytic or self-templating sets, groups of molecules that can catalyze or template their own formation, but these results require the prebiotic polymers to be of sufficient length, generally of at least 30 monomers.<sup>179,186</sup> Experimental work has demonstrated the formation of short peptide or nucleic acid multimers through step-growth mechanisms under prebiotic conditions,<sup>187–190</sup> but the statistical tendency of Flory-Schulz towards the formation of short chains imposes a significant barrier to the natural formation of chains of sufficient length to autocatalyze and self-template. Our results provide insight into a plausible mechanism by which this limitation could be overcome, in which initial oligomerization could produce reactant heterogeneity which biases the formation of longer chains. Such a polymerization driven length biasing away from Flory-Schulz behavior could have contributed to the initial formation of prebiotic polymers of length sufficient to autocatalyze, a crucial step on the pathway to biological life. Additionally, the sequence biasing and characteristic block length emergence we observe under similar conditions hints at the possibility of this same mechanism contributing to sequence templating and reproduction behaviors also critical to the functioning of life as it has subsequently evolved.

## 6.3 Future work

### 6.3.1 Experimental exploration of simulation results

The results presented in this work have demonstrated, *in silico*, that emergent assembly behaviors can alter the reaction kinetics and products of a polymerization process in ways not predicted by classical theories based on static reaction constants and bulk concentrations. However, several important questions need to be addressed regarding the translation of these results into *ex silico* copolymerizations. Can we observe, in an experimental copolymerization under appropriate conditions, polymerization-driven collective reactant behaviors that alter the solution environment and thereby alter polymer reaction kinetics? Can such polymerization-induced heterogeneities be employed experimentally to selectively bias copolymer sequences? How does emergent copolymer sequence biasing impact the dispersity, morphology, and phase behavior of the end reaction products? To address each of these questions, future work could utilize *in situ* sequence characterization techniques in a flow-based polymerization process to observe the sequence development within a model copolymerization at multiple stages of the reaction. Such an approach could determine the influence of reactant assembly and emergent heterogeneity on copolymer sequences and thereby test predictions made in the computational studies of this work.<sup>86,87</sup>

In order to better understand sequence development in copolymerization and to test the predictions from our computational and theoretical models, future work may employ NMR and LC-MS/MS sequence characterization techniques, alongside a combination of standard optical and chromatography diagnostic techniques to obtain kinetic and molecular weight information of a model system.<sup>5,191–193</sup> Such an approach could be further amplified by the use of flow-based polymerization setups, with on flow-line instrumentation allowing the *in situ* application<sup>191,193</sup> of the proposed characterization methods at various stages of the reaction in a straightforward manner through flow-fractioning. Such an approach would uniquely allow for the facile sampling of the reaction products at various reaction coordinates and monomer conversions, through fractioning off of the products at different flow times, as well as providing greater control over solution composition and reaction temperature.<sup>194</sup> Characterization of sequence development throughout the reactions could be performed using *ex situ* NMR capabilities for the identification of diad, triad, and pentad sequence fractions,<sup>55</sup> as well as LC-MS/MS characterization of flow-fractioned sample, utilizing the computational OLIGOSS method for ensemble sequencing.<sup>76</sup> The resulting sequence fractions, molecular weight distributions, and polymerization kinetics can

be compared directly to our molecular simulations to validate our model and obtain reactant level information on morphology and assembly during polymerization.

Additional possibilities for direct connection of the results in this work to experiment lie in the prediction of aggregate formation, composition, structure and ordering, as discussed in Chapters 4&5. The scale of these aggregates are more readily accessible to experimental measure than the species differentiation at the individual monomer required for sequence determination.<sup>5,14,55</sup> The formation of aggregates of sufficient size could be identified through a variety of different methods, with connections made to simulation results to support the model description or to refine the parameter space to more accurately reflect a particular system of interest. Static light scattering (SLS) and dynamic light scattering (DLS) are commonly used in tandem as a non-destructive approach, allowing for the determination of the scattering structure factor, which may be related to the radial distribution function calculated from simulation, as well as in the determining the effective hydrodynamic radius of polymeric aggregates.<sup>5,195,196</sup> Structural information of aggregates may also be investigated via a labeling approach, through the introduction of fluorescent<sup>197,198</sup> or electron spin<sup>199</sup> tags to the polymers themselves. Fluorescent labels could provide access to Förster resonance energy transfer (FRET) or fluorescence correlation spectroscopy (FCS), techniques commonly employed in the study of the aggregation and binding of biopolymers and which show promise in similar explorations for non-biological polymers.<sup>200,201</sup> Very recent work has even shown the potential for fluorescence coupled polymerization techniques to explore sequence information, albeit under significant constraint to the polymerization chemistry involved.<sup>202</sup> Similarly, the introduction of spin-tagged monomers would allow for electron paramagnetic resonance (EPR) measurements, which have been used to study polymer aggregation,<sup>203,204</sup> local structure and heterogeneity,<sup>205,206</sup> and even chain mobility within interfaces and aggregates.<sup>207,208</sup>

A potential candidate reaction for the explorations described is shown in Fig. 6.1, as adapted from work in Ref. [209]. This polymerization reaction relies on an alkene thiol “click chemistry” approach, and has been shown<sup>209</sup> to proceed in a step-growth fashion. Further, the reaction is photoinitiated, which would allow thorough mixing of reactants prior to initiation, and adapts well to flow-polymerization approaches, allowing for simultaneous initiation of the reaction and well-controlled reaction times via controlling dwell time in the photo-excited region of the reactor.<sup>172,194</sup> Modification of reactant affinities could be performed by appropriate selection of R groups for each of the monomer species, alongside appropriate solvent selection. For example, as the reaction is in DMSO, choosing a sulfonamide for the more solvophilic<sup>210</sup> species and a non-polar group such as a

phenyl group for the more solvophobic species, would allow for the type of selective solvent interactions we explore in Chapter 4. Further, these interactions could be adjusted by blending the solvent with either a less polar or more polar solvent than DMSO<sup>210</sup> to tune the effective attractive interactions experienced by the monomer species.

The simulations performed in this work, examining sequence biasing and kinetic implications of emergent assembly behaviors over a range of plausible non-bonded interaction strengths, comonomer reactivities, and backbone rigidities, could therefore be used to guide the design and execution of proposed experiments.<sup>58,86,87</sup> To facilitate direct comparisons between the experimental and computational work, new simulations may be performed under parameters corresponding to the synthetic conditions employed.

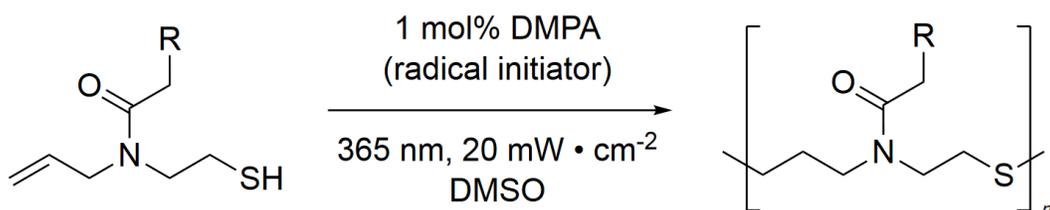


Figure 6.1: **Potential experimental reaction scheme.** A potential reaction scheme for future experimental exploration of results presented in this work is shown. Comonomer species would utilize the same thiol-allylamine backbone, with differing R groups. Monomers are photopolymerized by 365 nm light in the presence of DMPA (2,2-Dimethoxy-2-phenylacetophenone) photoinitiator. Adapted from a polymerization scheme presented in Ref. [209].

### 6.3.2 Simulation of additional polymerization schemes

While this work has focused on step-growth polymerization methods, many polymerization schemes of interest instead follow chain-growth kinetics.<sup>5</sup> Of particular interest are a family of radical based polymerization schemes, known as controlled radical polymerizations<sup>211,212</sup> (CRP), which include such commonly employed reaction methods as atom transfer radical polymerization<sup>26,213</sup> (ATRP), reversible addition fragmentation chain transfer (RAFT) polymerization,<sup>214</sup> and nitroxide-mediated polymerization<sup>215</sup> (NMP). While the precise nature of radical initiation and propagation mechanisms of each of these methods differ, each relies on generating a radical on growing chain through which the reaction

proceeds, and crucially, provides a mechanism for these radicals to be transferred between so-called “living” and “dormant” chains. This has the advantage of causing propagation to proceed more evenly throughout all chains in the system, reducing the chain-length and molecular weight dispersity of the end product. Further, there has been significant interest in using CRP from functionalized surfaces in “surface initiated” polymerizations<sup>216–219</sup> to generate functionalized surfaces and interfaces. The behavior of polymer-nanoparticle hybrid materials generated by such approaches have shown material properties of significant interest, and have been identified as having a number of open theoretical questions related to their solvation and self-assembly.<sup>219–222</sup> Such systems have attracted computational studies using both Monte Carlo<sup>80,223–225</sup> and molecular dynamics<sup>78,79</sup> approaches, but questions of sequence development and the impact of sequence on material properties in these systems are less explored.

Future work could utilize an adaptation of the model presented in this work to explore sequence development and self-assembly behaviors for CRP polymerization schemes, both in solution and in bulk or nanoparticle surface initiated systems. The author and coworkers have begun such adaptation with an eye towards answering a number of research questions. Analogous to the step-growth observations in this work, do non-bonded interactions, chain stiffness, and collective monomer behaviors impact the sequence and composition of a copolymer generated via CRP in a fashion not predicted by standard Mayo-Lewis<sup>5,74</sup> theory? Recent simulation work<sup>63,226</sup> has suggested breakdowns in Mayo-Lewis behavior in some CRP chemistries from analogous non-bonded interaction mechanisms, and these behaviors could be explored with more realistic and detailed dynamics with our model. Further, how might such behaviors differ in the presence of a surface, and do the geometric properties of the surface such as curvature, faceting, or grafted initiator density change these behaviors? A discussion of the preliminary work towards adapting the model herein to exploring such systems and answering these questions is presented as follows.

**Modeling a chain-growth, controlled radical polymerization.** The primary structure of the model follows closely with that of the step-growth copolymer work, with monomer units represented as three-particle beads, with a central particle and two linking particles all connected by intramonomer bonds represented via angular and linear harmonic potentials. Our standard **A** and **B** monomer types are included, along with new types of particles which represent the nanoparticle or surface, and the graft sites from which polymerization can be initiated. Graft sites are similarly represented as a three particle beads in an equivalent fashion to monomer beads, while nanoparticles are represented as single atoms. The nanoparticle is anchored at the origin, and nanoparticle beads are frozen in place, but

can interact with other particles. Graft beads placed randomly on the surface, and interact with the surface via a deep well Morse potential which affixes the graft site close to the surface but allows for dynamics at the point chains are grafted to the surface. Just as in our step-growth simulations, all unreacted monomers are randomly dispersed in the remaining volume, and all bonded and non-bonded interactions between graft beads and monomer beads centers are governed through the same interactions utilized in the step-growth work (see Ch. 2).

For simplicity, in modeling our CRP reaction scheme, we do not explicitly model the initiator agent, such as a photoinitiator or some other radical generating agent, and instead presume that the initiation step is sufficiently faster than the propagation step that initiation occurs approximately simultaneously in our system. Similarly, we do not explicitly model the chain-transfer agent, *i.e.*, the RAFT agent or ATRP catalyst which passes active radicals between chains. Instead, we represent this important feature of CRP schemes by rapidly switching sites where polymerization can occur, namely the graft bead ends or polymer chain ends, between the active and inactive states. This switching is done at random every thousand simulation timesteps,  $10^3 dt$ , which is very fast with respect to the actual reaction rate in our simulation. That radical initiation and transport occurs much more rapidly than propagation is a key feature of CRP, helping to ensure chains are more likely to grow at the same rate. With such a model, the fraction of living versus dormant chains may be chosen in a straightforward fashion, allowing for the exploration of the impact of various living chain fractions on the reaction behavior. Further development could consider the relative accessibility of a particular site to a chain transfer agent. This could entail assigning a weighted probability to sites being made active or inactive based on some sort of local neighboring criteria, or on the identity of the central bead of the active site, *i.e.*, an **A** monomer, **B** monomer, or graft site.

With the location of active radicals, and therefore living chains, determined as indicated above, the propagation step occurs in a similar fashion to our step-growth model. When a linking particle of an unbound monomer comes within  $0.2\sigma$  of an active site, a new bond is formed, and the radical propagates down the chain, with the end of the newly bound monomer becoming the new active site.

**Preliminary results of a chain-growth copolymerization model.** To examine the efficacy of the model described above in capturing features of a chain-growth copolymerization, as well as consider the potential of the model to explore the questions we propose, a small set of simulations were performed and analyzed. Fig. 6.2 shows preliminary results obtained from one such simulation trial, in which monomer attractions were set to

$\epsilon_{AA, BB} = k_B T$  and  $\epsilon_{AB} = 0$ , with chain stiffness set to a semiflexible value of  $l_p = 9.9$  (recall the step-growth cases explored in this work were  $l_p = 3.5$  for flexible chains and  $l_p = 16.5$  for stiff chains). Initiator concentration was set to  $f_I = 0.1$  as a fraction of initial monomer concentration, and the fraction of living initiators or chains was  $f_{\text{living}} = 0.2$ . Notably, the resultant chain length distribution (Fig. 6.2a) closely matches that of a Poisson distribution, with mean,  $\lambda$ , equal to the observed average degree of polymerization,  $\overline{X}_n$ . This result matches the predicted chain length distribution for a polymerization which proceeds with chain growth kinetics.<sup>5,73</sup> Further, in Fig. 6.2c, we observe that a semi-log plot of the degree of polymerization,  $\overline{X}_n$ , is linear in reaction time, suggesting that we are capturing first-order reaction kinetics expected for the propagation step of a chain-growth radical method.<sup>5,73</sup> Importantly, the chain length dispersity,  $\mathcal{D}$  remains within the bound of  $\mathcal{D} \lesssim 1.1$ , a value commonly identified as being a threshold for a “well controlled” radical polymerization.<sup>225</sup> Further, it both approaches and remains bounded by the theoretical prediction<sup>73,227</sup> of  $\mathcal{D} \approx 1 + \overline{X}_n / (1 + \overline{X}_n)^2$ , an expected dispersity for CRPs at high degree of polymerization. Taken together, these results suggest that our model successfully captures key kinetic features of a chain growth mechanism expected for a CRP scheme. Interestingly, we do not note the emergent increase to reaction rate or a breakdown of predicted chain length statistics that we observe in the step-growth case, though we do see the impacts of the attractive interactions on the sequence neighbor probabilities shown in Fig. 6.2d. Though this may be due to the fundamental difference in the reaction schema, given the very early occurrence of phase separation with respect to the reaction progression, it is also possible that these impacts could still be observed in other regimes. Future work would require careful consideration and refinement of relevant system timescales, attraction strengths, and system density in order to make this determination.

Another observation of interest in early explorations of this model is the impact of surfaces on the composition of the resultant polymers, shown in Fig. 6.3. These preliminary results show surface dependent changes to copolymer composition at the same monomer feed ratio, potentially indicating deviations from Mayo-Lewis theory during surface-initiated radical copolymerizations. The addition of surface interactions and the associated complexities of surface size, geometry, faceting, and grafting density create a wealth of opportunities for further study on the impact of these variables on sequence development and reaction kinetics in a CRP process. Further study is needed to understand this interplay, in order to both better predict the outcome of surface initiated CRP processes, and to potentially exploit behaviors of interest in informed materials design.

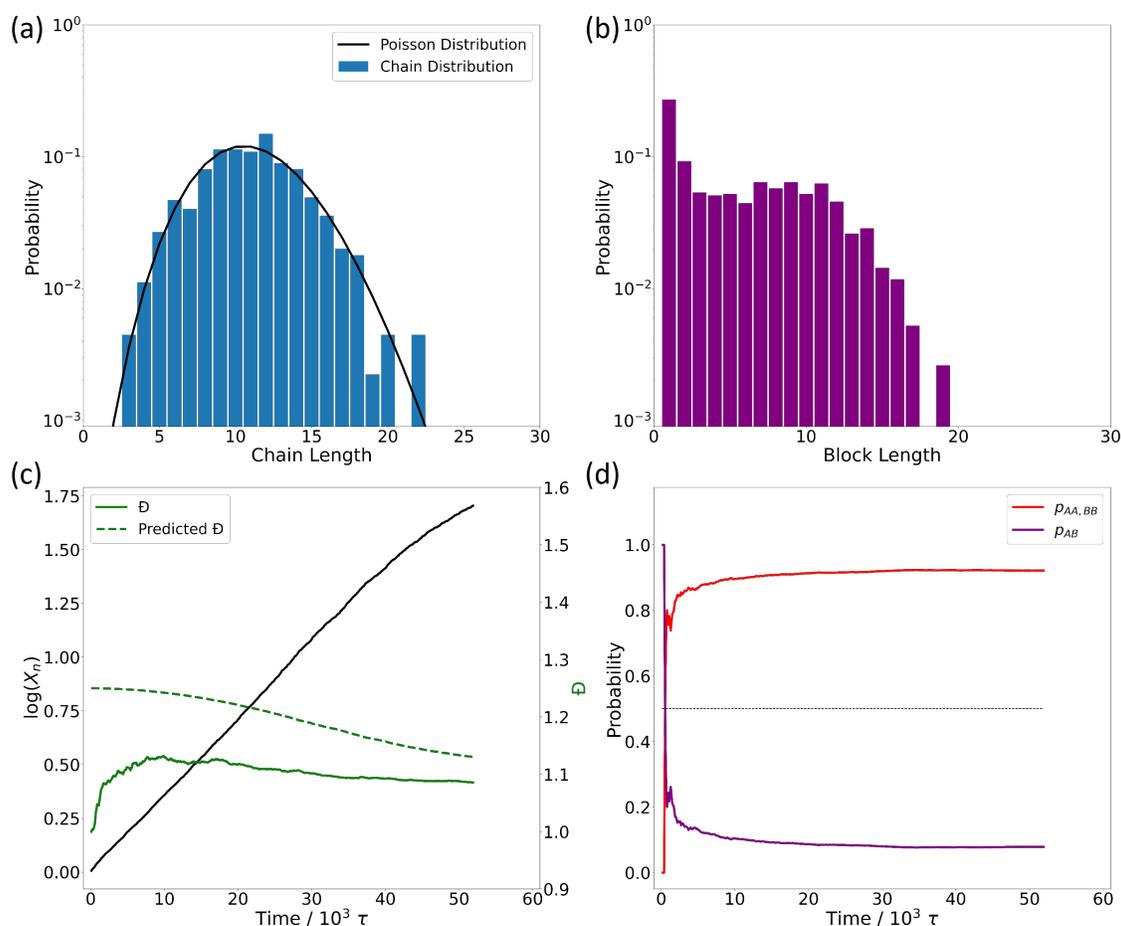


Figure 6.2: **Preliminary results for a single CRP simulation trial.** (a) Chain length distribution observed at 90% monomer conversion. The chain length distribution is plotted alongside a Poisson distribution with a mean,  $\lambda$ , equal to the observed average degree of polymerization,  $X_n$ . (b) The block length distribution observed at 90% monomer conversion. (c) The natural log of the average degree of polymerization,  $X_n$  (black line), as a function of simulation time. Also shown is the observed chain length dispersity,  $\mathcal{D}$  (solid green line), as a function of simulation time. The theoretical dispersity (dashed green line) is shown as predicted from the observed  $X_n$  at the simulation time indicated. (d) Sequence neighbor probabilities  $p_{AA,BB}$  (red line) and  $p_{AB}$  (purple line), as a function of simulation time. The black dashed line shows the value expected for a random sequence,  $p_{AA,BB} = p_{AB} = 0.5$ .

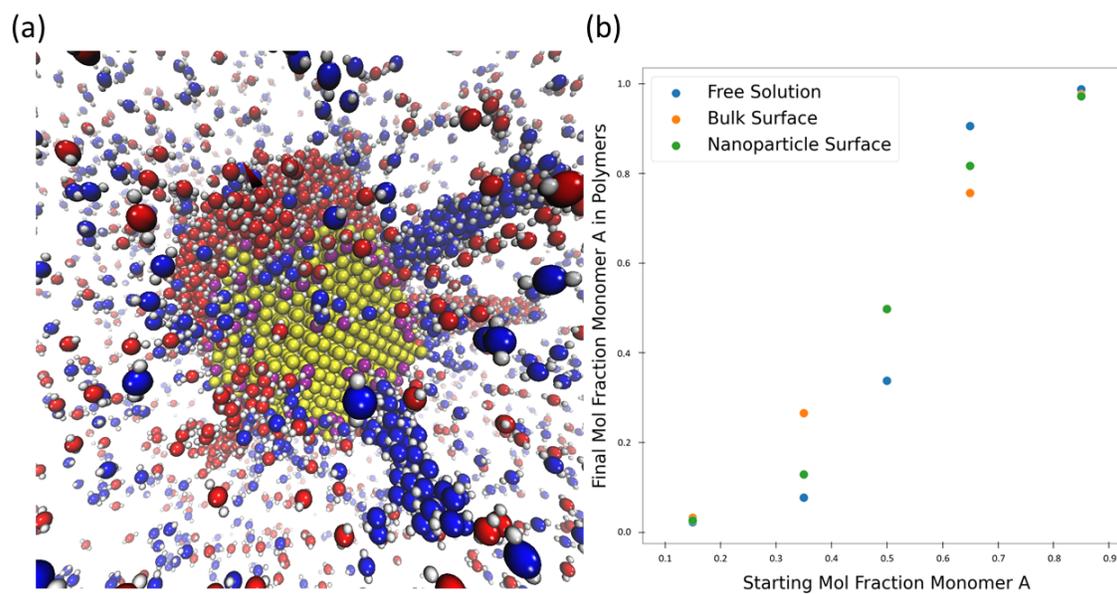


Figure 6.3: **Surface initiated CRP simulations.** (a) Visualization of a controlled radical copolymerization initiated from an icosahedron nanoparticle surface. Yellow spheres represent the nanoparticle, purple the surface initiators, red and blue the monomer types, and white the reactive end groups. (b)

## 6.4 Closing remarks

Synthetic polymers form the backbone of countless materials and are ubiquitous in scientific and engineering developments of the last century. This versatility arises, in part, from the ability to tailor polymers for specific applications, through informed design strategies which exploit chemical modification and structural variations. Control over the primary monomer sequence would open a vast space for further design refinement by exploiting the intrinsic connection between sequence, morphology, and bulk material properties. This potential for bespoke molecular design could lead to synthetic polymeric materials with improved mechanical strength, enhanced conductivity, unique optical characteristics, and even biomimetic assembly and information storage.

In order to harness this vast potential, further work is needed to advance the understanding of sequence development in polymerizations processes. Particular focus is needed on ways in which environmental factors in a reaction can control or bias sequence in a chemically agnostic fashion. The reactant heterogeneity, self-assembly, and self-templating behaviors we discuss herein apply to a broad class of polymerization processes, and the insight we have gleaned represents further progress towards unlocking generalized sequence control methods for synthetic polymer design and synthesis.

# Appendix A

## Chapter 3 Supplemental Details and Results

### A.1 Simulation details

The model we use here was described and discussed in detail in Chapter 2. All simulations were performed in LAMMPS.<sup>228</sup> Additional details specific to Chapter 3 are described below.

**Dimensionless units** All simulations used dimensionless LJ units within LAMMPS and are defined in terms of a chosen mass,  $m$ , length,  $\sigma$ , and energy,  $\epsilon$ . From these parameters, a characteristic time scale  $\tau \equiv \sqrt{\frac{m\sigma^2}{\epsilon}}$  is defined. In this work, we take  $m = 200$  Da,  $\sigma = 5$  Å, and  $\epsilon = 100 \text{ K} \cdot k_{\text{B}} = 1.38 \times 10^{-21}$  J, resulting in  $\tau = 7.75$  ps. For every monomer, the two linking particles have a mass of  $0.25 m$  and the central particle has a mass of  $0.5 m$ . The total monomer mass is then  $1.0 m$ , which is roughly the mass of a benzene ring with a 9-carbon alkyl substituent. The temperature for all simulations was set to 300 K. All simulations were run within a  $50\sigma \times 50\sigma \times 50\sigma$  cubic box with periodic boundary conditions. Each simulation contained 7200 monomers, yielding an overall density of  $\rho = 0.0576 \sigma^{-3}$ .

**Angular potential** Intramolecular and intermolecular angles are governed via the harmonic potential:

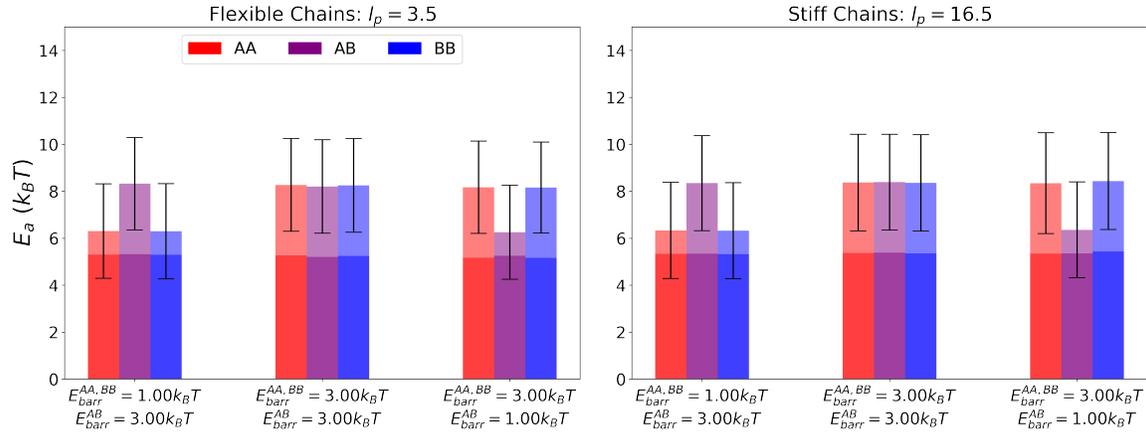
$$E_{\text{angle}}(\theta_{ijk}) = K_{ijk}^{\text{angle}} (\theta_{ijk} - \theta_0)^2,$$

in which  $\theta_{ijk}$  is the angle between particles  $i$ ,  $j$  and  $k$ ,  $\theta$  is the equilibrium angle and  $K_{ijk}^{\text{angle}}$  is the spring constant for the angle. For the intramonomer angle between particle **2-1-2'**,

$\theta_0 = 180^\circ$  and the spring constant is set to  $K_{212'}^{\text{angle}} = 5 \text{ } \epsilon \text{ rad}^{-1}$  and  $K_{212'}^{\text{angle}} = 50 \text{ } \epsilon \text{ rad}^{-1}$  for the flexible chains and stiff chains, respectively. For the intermonomer angle, **1-2-2'**,  $\theta_0 = 180^\circ$ ,  $K_{122'}^{\text{angle}} = 100 \text{ } \epsilon \text{ rad}^{-1}$  for both flexible chains and stiff chains.

### A.1.1 Activation energy

In Chapter 3, we explore a combination of activation energy such that  $\Delta E_a = E_a^{\text{AA, BB}} - E_a^{\text{AB}}$  ranges from  $-4 k_B T$  to  $4 k_B T$  in  $0.5 k_B T$  increments. Activation energies are composed of LJ repulsions between central particles, geometric constraints arising from intramonomer bonding between central and linking particles, and short-ranged repulsion between linking particles. Central LJ repulsions and geometric constraints are kept constant at a total value of  $5.3 k_B T$  for both flexible chains and stiff chains (Fig.A.1). The actual activation energy therefore will be equal to  $E_a = 5.3 + E_{\text{barr}}$ , where  $E_{\text{barr}}$  is the short-range repulsion between linking particles and is varied from 0 to  $4 k_B T$  as listed in Table A.1.



**Figure A.1: Measured activation energy.** The activation energies observed in bonding events sampled from simulations. Sample populations were drawn from all four attractions cases explored in this work, and are shown for three activation energy cases:  $E_{\text{barr}}^{\text{AA, BB}} = 3k_B T$ ,  $E_{\text{barr}}^{\text{AB}} = 1k_B T$ ;  $E_{\text{barr}}^{\text{AA, BB}} = E_{\text{barr}}^{\text{AB}} = 3k_B T$ ; and  $E_{\text{barr}}^{\text{AA, BB}} = 1k_B T$ ,  $E_{\text{barr}}^{\text{AB}} = 3k_B T$ . Sample populations are further separated by monomer pair, and both flexible (left) and stiff chain (right) results are shown. Mean values of the measured activation energies for each combination are shown, with the error bars representing the standard deviation. The dark shades of colors represent the constant component of activation energy, which is composed of LJ repulsions and geometric constraints, and has an average value of  $5.3 k_B T$  for both flexible and stiff chains. The lighter shaded region represents the short-ranged repulsion between type-2 particles from  $E_{\text{barr}}$ , which is varied across simulations as described in Table A.1 to control the total activation energy.

$E_{\text{barr}}^{\text{AA,BB}} (k_B \text{T})$	$E_{\text{barr}}^{\text{AB}} (k_B \text{T})$	$E_{\text{barr}}^{\text{AA,BB}} - E_{\text{barr}}^{\text{AB}} (k_B \text{T})$
4.0	0.0	4.0
3.5	0.0	3.5
3.0	0.0	3.0
3.0	0.5	2.5
3.0	1.0	2.0
2.0	0.0	2.0
3.0	1.5	1.5
2.0	0.5	1.5
3.0	2.0	1.0
2.0	1.0	1.0
3.0	2.5	0.5
2.0	1.5	0.5
3.0	3.0	0.0
2.0	2.0	0.0
2.5	3.0	-0.5
1.5	2.0	-0.5
2.0	3.0	-1.0
1.0	2.0	-1.0
1.5	3.0	-1.5
0.5	2.0	-1.5
1.0	3.0	-2.0
0.0	2.0	-2.0
0.5	3.0	-2.5
0.0	3.0	-3.0
0.0	3.5	-3.5
0.0	4.0	-4.0

**Table A.1: Short-range repulsion between linking particles.** Simulations were done with these values of  $E_{\text{barr}}^{\text{AA,BB}}$  and  $E_{\text{barr}}^{\text{AB}}$ , which define the height the the short-range repulsion (see Eq. 2 in the main text), as well as their difference  $E_{\text{barr}}^{\text{AA,BB}} - E_{\text{barr}}^{\text{AB}}$ . For all combination of activation energies, the strength of the LJ repulsion and the geometric constrains are kept constant irrespective of monomer pair identities.

## A.2 Polymerization kinetics and chain length distribution.

We examined the kinetics of the polymerization via the growth of degree of polymerization,  $X_n$ , and the dispersity,  $\mathcal{D}$ , throughout the polymerization. Flory-Schulz theory predicts a linear relationship between  $\mathcal{D}$  and  $p$  via the equation  $\mathcal{D} = 1 + p$  in ideal conditions.<sup>73</sup> As a result,  $\mathcal{D}$  approaches a plateau of  $\mathcal{D} = 2$  as the reaction proceeds. In Fig. A.2, we plot  $X_n$  over time (Fig. A.2, top row) and  $\mathcal{D}$  as a function of reaction extent  $p$  (Fig. A.2, bottom row). We observe non-linear growth of  $X_n$  over time in the case of higher attractions (Fig. A.2, a - d), suggesting altered rate behavior as the reaction proceeds, indicating a departure from kinetics governed by Flory's principle of equal reactivity.<sup>73</sup> Moreover, in these higher attraction cases we also observed non-linear growth of  $\mathcal{D}$ , reaching values  $\mathcal{D} > 2$  (Fig. A.2 e-h), indicating the increased presence of longer chains in the system as compared to that predicted under ideal polymerization kinetics.

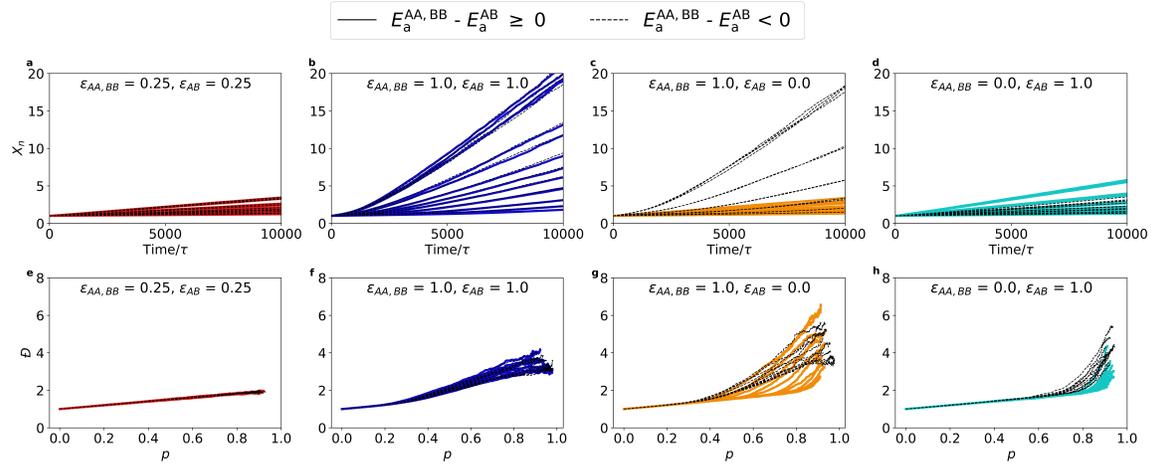


Figure A.2: **Polymerization kinetics including degree of polymerization  $X_n$  and Dispersity  $\mathcal{D}$ .**  $X_n$  (top row) and  $\mathcal{D}$  (bottom row) are shown for four different attraction cases. Solid lines represent cases of  $E_a^{AA, BB} - E_a^{AB} \geq 0$ , with colors corresponding to Fig. 3 in the main text, and black dashed lines represent cases of  $E_a^{AA, BB} - E_a^{AB} < 0$ . The results for  $X_n$  at  $\epsilon_{\text{all}} = 0.25 k_B T$  have been previously shown in Fig. 2(a) in the main text and are included here again for ease of comparison. The results plotted here were averaged over three independent simulation trials for each parameter set.

In Fig. A.3, we plot the chain length distributions for flexible chains. Under ideal kinetics, the Flory-Schulz distribution predicts the probability of finding a chain length  $n$  via

$$P_n = (1 - p)p^{n-1},$$

where  $p$  is the reaction extent. The expected Flory-Schulz distributions is plotted in each case with a black line. In the case of low inter-monomer attraction strengths, shown in the first column of Fig. A.3, the chain length distributions show good agreement with the predicted Flory-Schulz distribution. By contrast, other higher attraction cases deviate from the expected Flory-Schulz statistics. These observation are consistent with the non-ideal kinetics observed in Fig. A.2.

### A.3 Markovian block length distribution

In Markovian statistics, where the probability of having a monomer type in the chain depends solely on the identity of the preceding monomer, the expected probability of having of a continuous block of **A** or **B** in the copolymer sequence can be calculated via:

$$P(n) = (1 - pp_{AA,BB})(pp_{AA,BB})^{n-1},$$

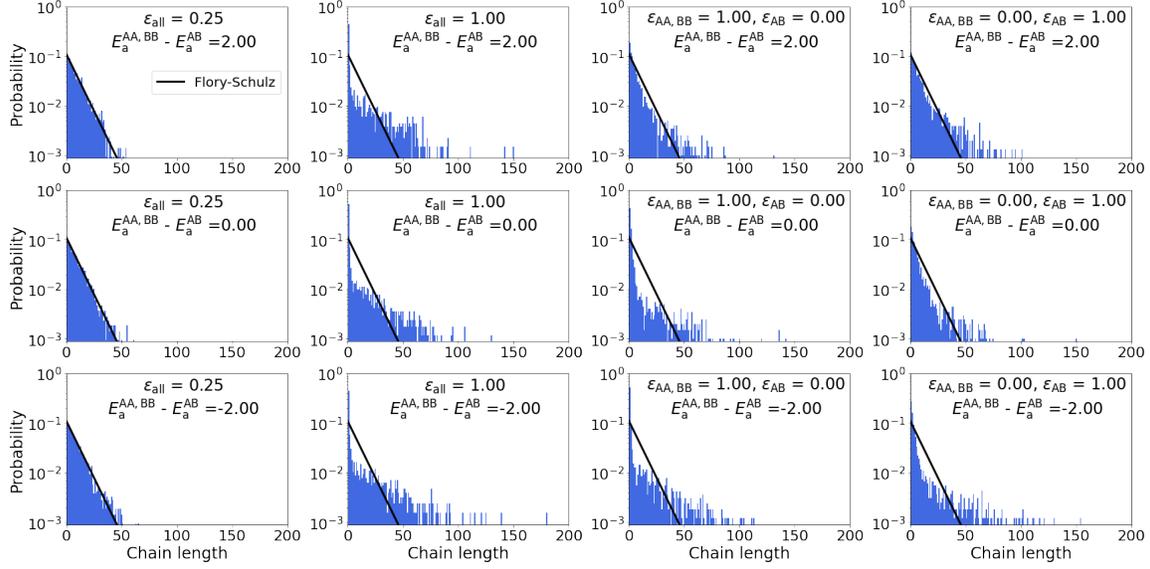
where  $p$  is the reaction extent and  $p_{AA,BB}$  is the probability of having a like nearest neighbor in the sequence. The block length distribution have been shown for the case of  $\epsilon_{AA,BB} = k_B T$ ,  $\epsilon_{AB} = 0$  in the main text of Chapter 3 and additional results for flexible chains are also shown here in Fig. A.4.

### A.4 Supplemental results for stiff chains

In Fig. A.5 and Fig. A.6 , we show additional chain length distributions and block length distributions for stiff chains. Note that, as previously observed in Ref. [59],  $E_a^{AA,BB} - E_a^{AB}$  are nearly identical for both flexible and stiff chains, therefore the bonding probability as predicted by equation Eq. 26 in the main text remains the same in each case. The chain length distributions for the higher attraction cases (Fig. A.5, columns 2-4) and the block length distributions in the case of  $1k_B T$  of attractions between only like monomers at  $\Delta E_a \geq 0$  (Fig. A.6, column 3) deviate from the ideal predictions.

### A.5 Selective solvent results

In addition, we have performed simulations in the condition that **A** and **B** are asymmetric so that  $\epsilon_{AA} \neq \epsilon_{BB}$  and  $E_a^{AA} \neq E_a^{BB}$ . In these cases, the bonding probability of having **A-A**



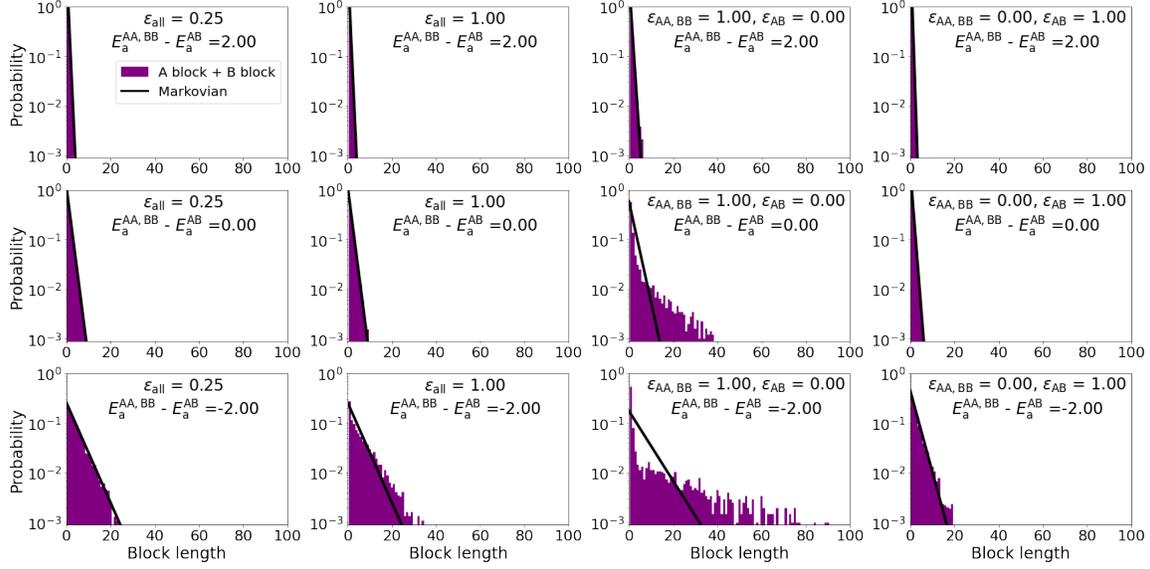
**Figure A.3: Chain-length distributions for flexible chains.** The chain length distributions (blue histograms) for four attraction cases and three combinations of activation energy. Columns represent each attraction strength combination as follows:  $\epsilon_{\text{all}} = 0.25 k_B T$  (first column);  $\epsilon_{\text{all}} = k_B T$  (second column);  $\epsilon_{\text{AA, BB}} = k_B T$ ,  $\epsilon_{\text{AA, BB}} = 0$  (third column); and  $\epsilon_{\text{AA, BB}} = 0$ ,  $\epsilon_{\text{AA, BB}} = k_B T$  (fourth column). Rows represent the activation energy combinations as follows:  $E_a^{\text{AA, BB}} = 8.3$ ,  $E_a^{\text{AB}} = 6.3 k_B T$  (top row);  $E_a^{\text{AA, BB}} = 8.3$ ,  $E_a^{\text{AB}} = 8.3 k_B T$  (middle row); and  $E_a^{\text{AA, BB}} = 6.3$ ,  $E_a^{\text{AB}} = 8.3 k_B T$  (bottom row). The black solid line shows the expected Flory-Schulz distribution in each case.

and **B-B** in the sequence is no longer equivalent.

Specifically, we ran simulations at  $\epsilon_{\text{AA}} = k_B T$ ,  $\epsilon_{\text{BB}} = 0.1 k_B T$ , and  $\epsilon_{\text{AB}} = \sqrt{\epsilon_{\text{AA}} \epsilon_{\text{BB}}} \approx 0.3 k_B T$  via Lorentz-Berthlot combining rules. The choice of LJ attractions is described in more detail in Ref. [86]. To explore the combined effects of activation energy and selective solvent conditions, we examined the system at  $E_a^{\text{AA}} > E_a^{\text{AB}} > E_a^{\text{BB}}$ ,  $E_a^{\text{AA}} = E_a^{\text{AB}} = E_a^{\text{BB}}$ , and  $E_a^{\text{AA}} < E_a^{\text{AB}} < E_a^{\text{BB}}$ .

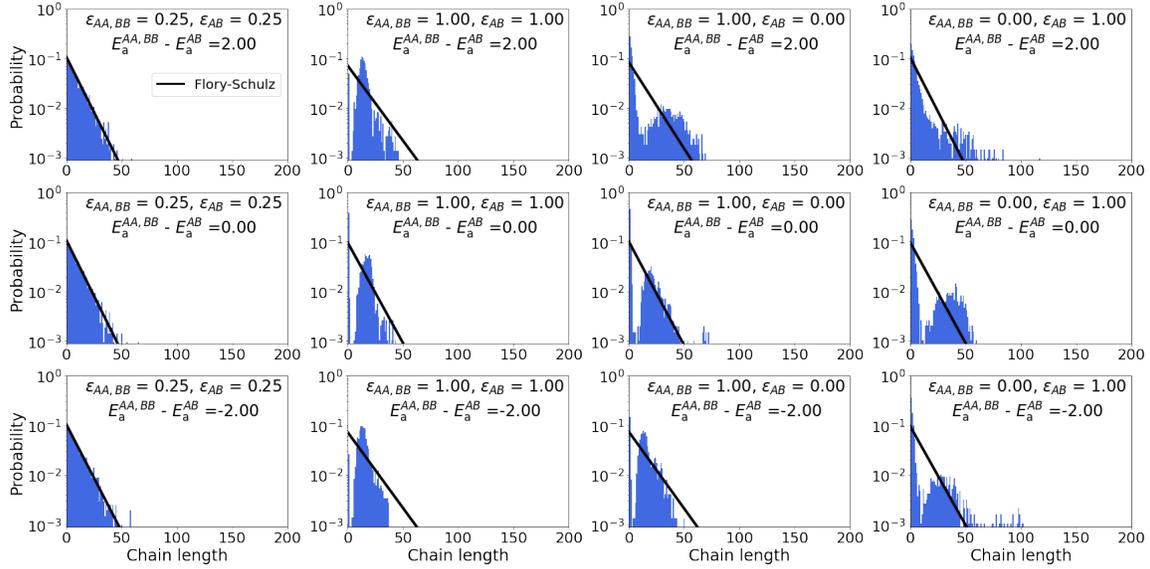
In Figure. A.7, we show the block length distributions separately for **A** blocks (in red) and **B** blocks (in blue) atop the chain length distributions (in grey shaded histogram). Sample resulting structures are also included on the top-right corner of the figures.

In all cases that are explored, the block length distributions of **B** agree with what is predicted from Markovian statistics while the block length length distributions of **A** deviate from predictions. Though the degree of deviation from Markovian predictions is similar in each case shown, due to the equivalent attraction strengths in each, the overall shift towards longer **A**-blocks depends on the relative activation energy of  $E_a^{\text{AA}}$  to  $E_a^{\text{AB}}$ ; reducing the en-

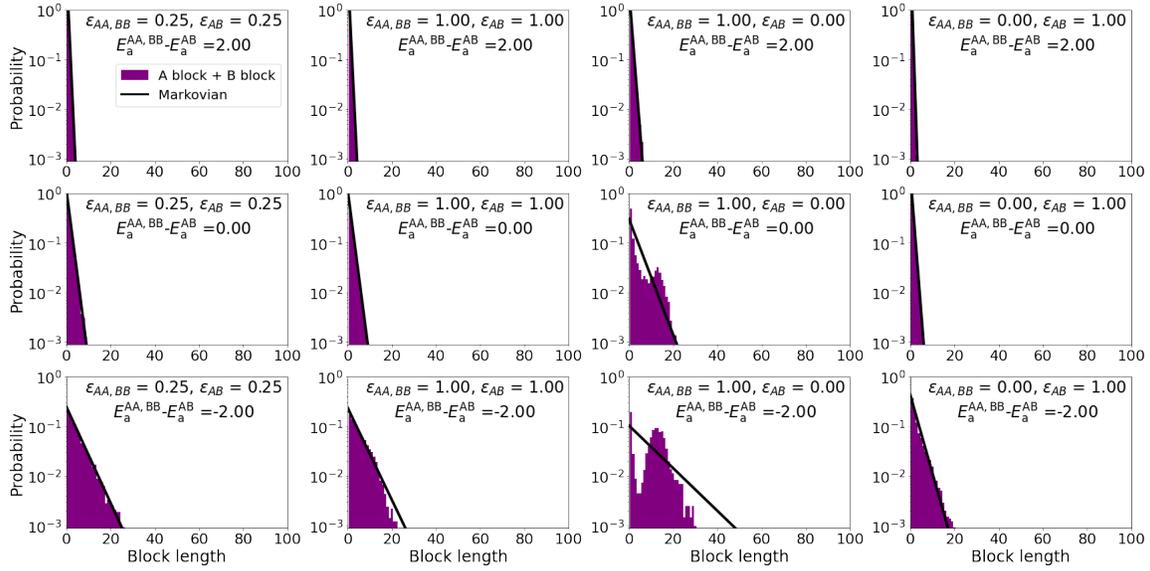


**Figure A.4: Block length distributions for flexible chains.** The block length distributions for **A** and **B** block are shown in purple histograms for four attraction cases and three combinations of activation energy. Columns represent each attraction strength combination as follows:  $\epsilon_{\text{all}} = 0.25 k_B T$  (first column);  $\epsilon_{\text{all}} = k_B T$  (second column);  $\epsilon_{\text{AA, BB}} = k_B T, \epsilon_{\text{AA, BB}} = 0$  (third column); and  $\epsilon_{\text{AA, BB}} = 0, \epsilon_{\text{AA, BB}} = k_B T$  (fourth column). Rows represent the activation energy combinations as follows:  $E_a^{\text{AA, BB}} = 8.3 k_B T, E_a^{\text{AB}} = 6.3 k_B T$  (top row);  $E_a^{\text{AA, BB}} = 8.3 k_B T, E_a^{\text{AB}} = 8.3 k_B T$  (mid row); and  $E_a^{\text{AA, BB}} = 6.3 k_B T, E_a^{\text{AB}} = 8.3 k_B T$  (bottom row). The black solid line shows the expected Markovian distribution in each case.

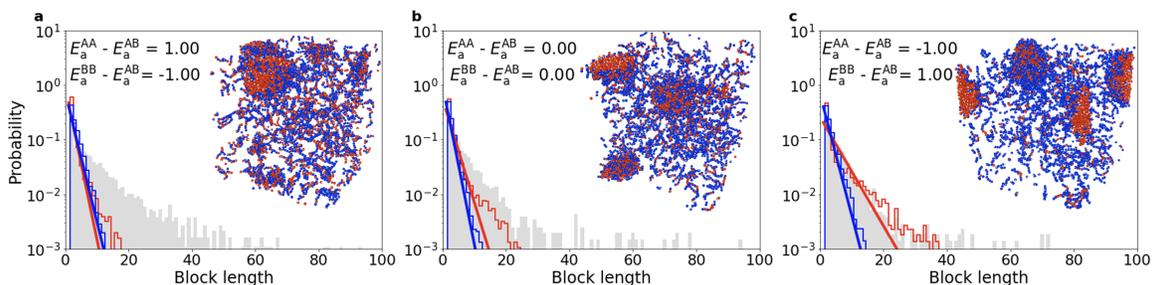
ergy barrier for **A–A** bond formation resulting in a shift towards greater lengths, and vice versa. From the corresponding structures that are included on the top-right corner of the figures, the effective self-attraction of monomer **A** are sufficient to drive their condensation into an **A**-rich phase, resulting in the shift in block length distributions from the Markovian statistics. In contrast, monomer **B** has a weaker effective self-attraction and remains dispersed. As a result, their block length distribution are still well described by the Markovian statistics.



**Figure A.5: Chain-length distributions for stiff chains.** The chain length distributions (blue histograms) for four attraction cases and three combinations of activation energy. Columns represent each attraction strength combination as follows:  $\epsilon_{\text{all}} = 0.25 k_B T$  (first column);  $\epsilon_{\text{all}} = k_B T$  (second column);  $\epsilon_{AA, BB} = k_B T, \epsilon_{AA, BB} = 0$  (third column); and  $\epsilon_{AA, BB} = 0, \epsilon_{AA, BB} = k_B T$  (fourth column). Rows represent the activation energy combinations as follows:  $E_a^{AA, BB} = 8.3, E_a^{AB} = 6.3 k_B T$  (top row);  $E_a^{AA, BB} = 8.3, E_a^{AB} = 8.3 k_B T$  (mid row); and  $E_a^{AA, BB} = 6.3, E_a^{AB} = 8.3 k_B T$  (bottom row). The black solid line shows the expected Flory-Schulz distribution in each case.



**Figure A.6: Block length distributions for stiff chains.** Block length distributions for flexible chains. The block length distributions for **A** and **B** block are shown in purple histograms for four attraction cases and three combinations of activation energy. Columns represent each attraction strength combination as follows:  $\epsilon_{\text{all}} = 0.25 k_B T$  (first column);  $\epsilon_{\text{all}} = k_B T$  (second column);  $\epsilon_{AA, BB} = k_B T, \epsilon_{AA, BB} = 0$  (third column); and  $\epsilon_{AA, BB} = 0, \epsilon_{AA, BB} = k_B T$  (fourth column). Rows represent the activation energy combinations as follows:  $E_a^{AA, BB} = 8.3 k_B T, E_a^{AB} = 6.3 k_B T$  (top row);  $E_a^{AA, BB} = 8.3 k_B T, E_a^{AB} = 8.3 k_B T$  (mid row); and  $E_a^{AA, BB} = 6.3 k_B T, E_a^{AB} = 8.3 k_B T$  (bottom row). The black solid line shows the expected Markovian distribution in each case.



**Figure A.7: Block length distributions for selective solvent conditions.** Block length distributions observed under attraction strengths modeling selective solvent conditions where  $\varepsilon_{AA} = k_B T$ ,  $\varepsilon_{AB} = 0.3k_B T$ , and  $\varepsilon_{BB} = 0.1k_B T$ . The block length distributions for **A** blocks (shown in red) and **B** blocks (shown in blue) are presented separately. Three combinations of activation energy are shown:  $E_a^{AA} = 8.3 k_B T$ ,  $E_a^{AB} = 7.3 k_B T$ ,  $E_a^{BB} = 6.3 k_B T$  (left column);  $E_a^{AA} = 7.3 k_B T$ ,  $E_a^{AB} = 7.3 k_B T$ ,  $E_a^{BB} = 7.3 k_B T$  (middle column); and  $E_a^{AA} = 6.3 k_B T$ ,  $E_a^{AB} = 7.3 k_B T$ ,  $E_a^{BB} = 8.3 k_B T$  (right column). The solid red and blue lines represent the predicted Markovian statistics for **A** and **B** blocks, respectively, as calculated from the observed bonding probabilities of  $p_{AA}$  and  $p_{BB}$  and the reaction extent  $p$ . For each case, a representative system structure, taken at  $p = 0.9$ , is inset, with **A** monomers shown in red and **B** monomers shown in blue.

# Appendix B

## Chapter 4 Supplemental Details and Results

### B.1 Simulation details

The model we use here was described and discussed in detail in Chapter 2. All simulations were performed in LAMMPS.<sup>228</sup> Additional details specific to Chapter 4 are described below.

**Dimensionless units** All simulations used dimensionless LJ units within LAMMPS and are defined in terms of a chosen mass,  $m$ , length,  $\sigma$ , and energy,  $\epsilon$ . From these parameters, a characteristic time scale  $\tau \equiv \sqrt{\frac{m\sigma^2}{\epsilon}}$  is defined. In this work, we take  $m = 200$  Da,  $\sigma = 5$  Å, and  $\epsilon = 100 \text{ K} \cdot k_{\text{B}} = 1.38 \times 10^{-21}$  J, resulting in  $\tau = 7.75$  ps. For every monomer, the two linking particles have a mass of  $0.25 m$  and the central particle has a mass of  $0.5 m$ . The total monomer mass is then  $1.0 m$ , which is roughly the mass of a benzene ring with a 9-carbon alkyl substituent. The temperature for all simulations was set to 300 K. All simulations were run within a  $50\sigma \times 50\sigma \times 50\sigma$  cubic box with periodic boundary conditions. Each simulation contained 7200 monomers, yielding an overall density of  $\rho = 0.0576 \sigma^{-3}$ .

**Persistence length and the angular potential** Intramolecular angles are governed by the harmonic potential:

$$E_{\text{angle}}(\theta_{ijk}) = K_{ijk}^{\text{angle}} (\theta_{ijk} - \theta_0)^2, \quad (\text{B.1})$$

where  $\theta_{ijk}$  is the angle between particles  $i$ ,  $j$ , and  $k$ ,  $\theta_0$  is the value at the minimum of the angular potential, and the constant,  $K_{ijk}^{\text{angle}}$ , is the spring constant for that angle. For the

intramonomer angle, **2-1-2**,  $\theta_0 = 180^\circ$  and  $K_{212}^{\text{angle}}$  was set to  $5 \text{ } \epsilon \cdot \text{rad}^{-2}$  for flexible chains and  $50 \text{ } \epsilon \cdot \text{rad}^{-2}$  for stiff chains. Previous work with this model has determined that these values correspond to persistence lengths of  $l_p = 3.5$  for flexible chains and  $l_p = 16.5$  for stiff chains.<sup>59</sup> For the intermonomer angle, **1-2-2'**,  $\theta_0 = 180^\circ$  and  $K_{122'}^{\text{angle}} = 100 \text{ } \epsilon \cdot \text{rad}^{-2}$  in all cases.

**Langevin dynamics** Langevin dynamics was implemented in LAMMPS using the “fix langevin” command. A *damp* parameter is used to control the diffusion rate and the relaxation rate of the temperature, which is discussed in detail in the Supplemental Information of Ref. [3]. In the simulations in this work, the *damp* parameter is set to  $0.1 \tau$ , resulting in a viscosity of  $\gamma = 0.1 \text{ mPa} \cdot \text{s}$ . For further discussion of this parameter, see Ref. [59].

### B.1.1 Attractive interaction strengths and Lorentz-Berthelot combining

After choosing  $\epsilon_{AA}$  and  $\epsilon_{BB}$  as described in Methods, we utilized Lorentz-Berthelot combining rules<sup>92</sup> to define an attraction strength between our more and less solvophobic monomers, **A** and **B**, respectively, setting  $\epsilon_{AB} = \sqrt{\epsilon_{AA}\epsilon_{BB}}$ . As we are utilizing dimensionless, reduced unit values for  $\sigma_{AA}$  and  $\sigma_{BB}$ , this expression is also equivalent to Kong combining rules.<sup>91</sup> Table B.1 shows all combinations of the  $\epsilon_{\text{att}}$  values explored in this work.

### B.1.2 Activation energies

To ensure that having a fixed value for  $\epsilon_{\text{rep}}$ , plus the additional soft repulsive potential between type **2** particles, was maintaining a consistent activation energy across monomer type pairs for our range of simulation parameters, we tabulated the activation energies for successful bond formations for a set of  $\sim 52000$  bonding events drawn from the full range of simulation parameters explored in this work. Activation energies are calculated from the total potential energy of the two bonding monomers arising from the LJ potential between type **1** particles, the soft repulsive potential between type **2** particles, and the intramonomer bond lengths and angles. The values of each of these contributions above their minimum energy are summed at the time of bond formation to obtain the total activation energy for the bonding event. Figure B.1 shows the results of this comparison, broken down by monomer type pairs, demonstrating consistent activation energies across monomer type pairing.

$\epsilon_{AA} (k_B T)$	$\epsilon_{AB} (k_B T)$	$\epsilon_{BB} (k_B T)$
1.0	0.32	0.1
1.0	0.55	0.3
1.0	0.71	0.5
0.75	0.27	0.1
0.75	0.47	0.3
0.75	0.61	0.5
0.5	0.22	0.1
0.5	0.39	0.3
0.5	0.5	0.5

Table B.1: **Lennard-Jones attractions.** Columns show  $\epsilon_{AA}$ ,  $\epsilon_{AB}$ , and  $\epsilon_{BB}$ , which are the well-depths for the attractive portion of the LJ interactions (see Eq. 2.1 in the main text of Chapter 2) for **A** to **A**, **A** to **B** and **B** to **B** monomers, respectively. Values for  $\epsilon_{AA}$  and  $\epsilon_{BB}$  were chosen to reflect a copolymerization in which monomer species **A** is more solvophobic than species **B**. From the values of  $\epsilon_{AA}$  and  $\epsilon_{BB}$ ,  $\epsilon_{AB}$  values were calculated according to Lorentz-Berthelot rules.<sup>92</sup> For all combination of attraction strengths, the strength of repulsive interactions,  $\epsilon_{\text{rep}}$ , is kept at a constant value of  $1.33 k_B T$ , irrespective of monomer identity.

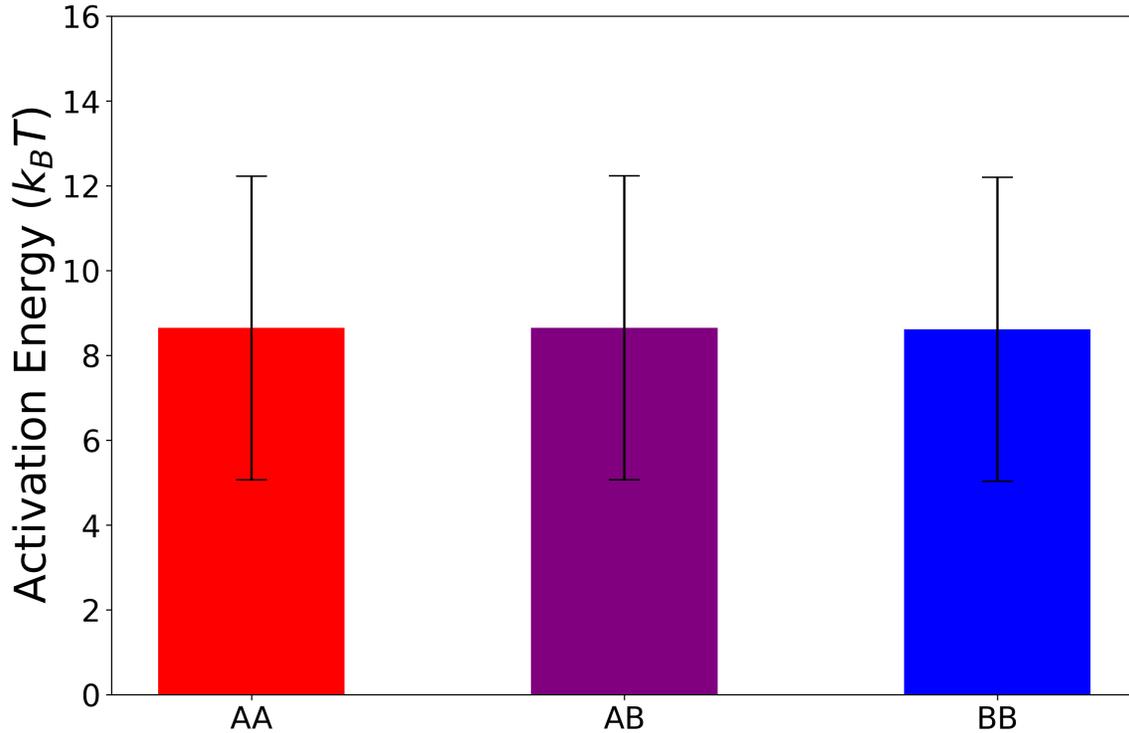


Figure B.1: **Observed activation Energies.** The average activation energies for bond formation events of **A** to **A**, **A** to **B**, and **B** to **B** monomer pairs for the full range of simulation parameters explored. Error bars represent the standard deviation in the recorded values for each monomer type pairing.

## B.2 Calculation of the first coordination number from $g(r)$

In liquids the first coordination number is commonly defined in terms of the radial distribution function,  $g(r)$ , and is taken as the number of atoms within the first coordination shell.<sup>107</sup> With this definition, we calculate the first coordination number for a given monomer species combination as:

$$n_1^{ij} = 4\pi\rho \int_0^{r'} r^2 g_{ij}(r) dr, \quad (\text{B.2})$$

where  $\rho$  is the number density of the system,  $r$  is the radial distance from the center of a fixed reference particle, and  $r'$  is the value of  $r$  at which  $g_{ij}(r)$  reaches the first minimum

after the initial peak. Indices  $i$  and  $j$  correspond to the possible monomer species pairs within the system, namely **AA**, **AB**, and **BB**.

### B.3 Quantifying deviations from Markovian block statistics

To quantitatively explore the deviation from Markov statistics in our system, we compared the Markovian predictions for dispersity and block length distribution with the results of our simulations. We determined the mean standard error (MSE) between the observed and predicted block dispersities (Equation 2.17) and calculated the Wasserstein metric,<sup>96,97</sup>  $W_p$ , as a measure of the statistical distance between the Markovian distributions and our observed block length distributions. Fig. B.2 shows the results of each of these metrics, as a function of  $\Delta\epsilon$ . With the increase in attraction asymmetry between **A** and **B** species, characterized by  $\Delta\epsilon$ , the difference between the Markovian distribution and the observed distribution of **A** block lengths also increases. This deviation demonstrates the non-Markovian nature of **A**-block formation under conditions in which attractive interactions drive monomer self-assembly. Notably, however, the Markovian distribution is a better match to the observed distribution of **B** block lengths, as  $W_p$  for the **B** blocks is lower than that for the **A** blocks and also largely independent of  $\Delta\epsilon$ . Indeed, the dispersities and length distributions of the **B** blocks are still well described by Markovian chain statistics for each interaction strength combination explored in this work.

The predicted block dispersities from Equation 2.17 are utilized to examine the deviation from Markovian behavior in our system, as described in Section B.3 and shown in Fig. B.2a below.

### B.4 Local nematic order parameter for flexible chains

The local nematic order parameter was calculated as described in Chapter 2. Fig. B.3 shows the value of the order parameter,  $\overline{S_{\text{local}}}$ , for flexible chains throughout the reaction at each of the attractive interactions explored in this work. Flexible polymers do not show significant nematic ordering regardless of attraction strength. Thus, the peak value for  $\overline{S_{\text{local}}}$  observed in Fig. B.3 is used as a threshold value for nematic ordering in stiff chains, as shown in Fig. 4.4 and described in the main text of Chapter 4.

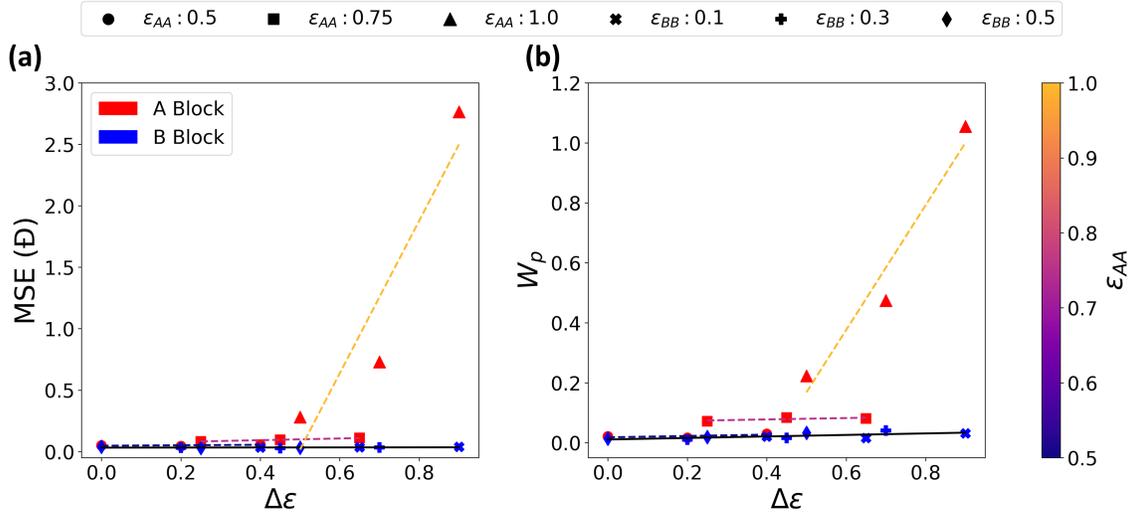


Figure B.2: **Error metrics for the predicted and observed block dispersities and block length distributions.** (a) The mean standard error (MSE) between the observed dispersity,  $\mathcal{D}$ , in block lengths, and the associated prediction for block length  $\mathcal{D}$  based on Markovian statistics. (b) The Wasserstein metric,  $W_p$ , a statistical distance between our observed block distribution and the Markovian distribution. In both (a) and (b), the block type, **A** or **B**, is shown by marker coloration. The marker shape indicates the value of attractions –  $\varepsilon_{AA}$  for **A** blocks and  $\varepsilon_{BB}$  for **B** blocks. Dashed lines show trends in the values for **A** blocks for cases with equal  $\varepsilon_{AA}$ , with the value of  $\varepsilon_{AA}$  indicated by line coloration. The black line is the trend in **B** block values for all simulation parameters. Observed distributions were collected from three independent trials for each simulation parameter set.

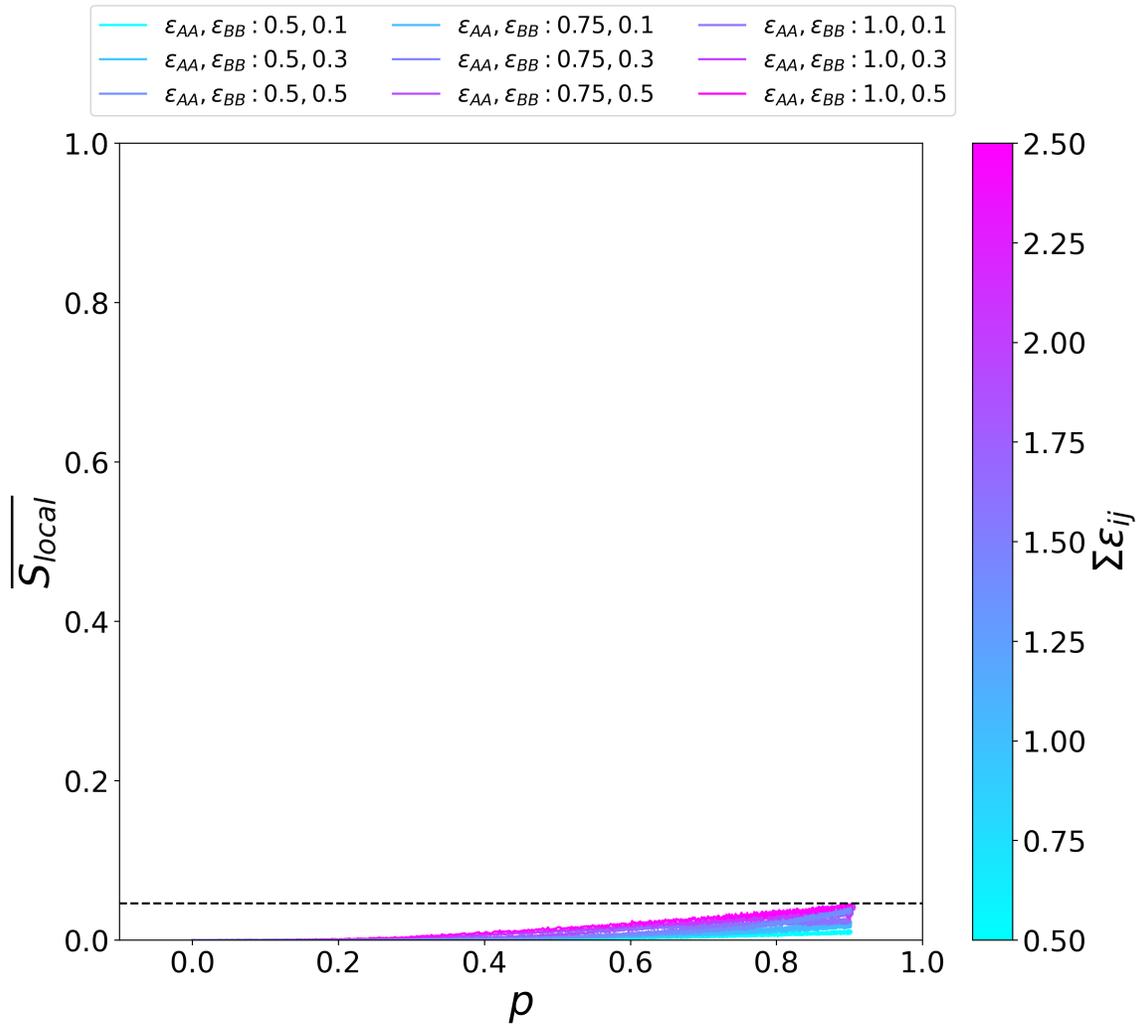


Figure B.3: **Local nematic order parameter for flexible chains.** Local nematic order parameter,  $\overline{S}_{local}$ , as a function of reaction extent,  $p$ , over the range of attraction strengths explored for flexible-chain,  $l_p = 3.5$ , polymers. The dashed horizontal line indicates the highest value for  $\overline{S}_{local}$  seen in this case. This line is reproduced in Fig. 4.4 in the main text of Chapter 4.

## B.5 Polymerized fraction

Asymmetric attractive interactions produce different rates of polymerization between monomer species. The fraction of each monomer type that has undergone polymerization depends on the balance of these interactions as well as the development of the emergent aggregate phase. We simulate a fixed quantity of monomers with an equal amount of **A** and **B**. Without a continuous reactant feed, the polymerized fractions of the monomer species are also impacted by the total availability of each monomer species. This impact is particularly notable late in the reaction, where the more rapidly polymerizing monomers near complete conversion. As such, the polymerized fractions of the two monomer species varies in a complex fashion throughout the reaction. This variation can be seen in Fig. B.4 below.

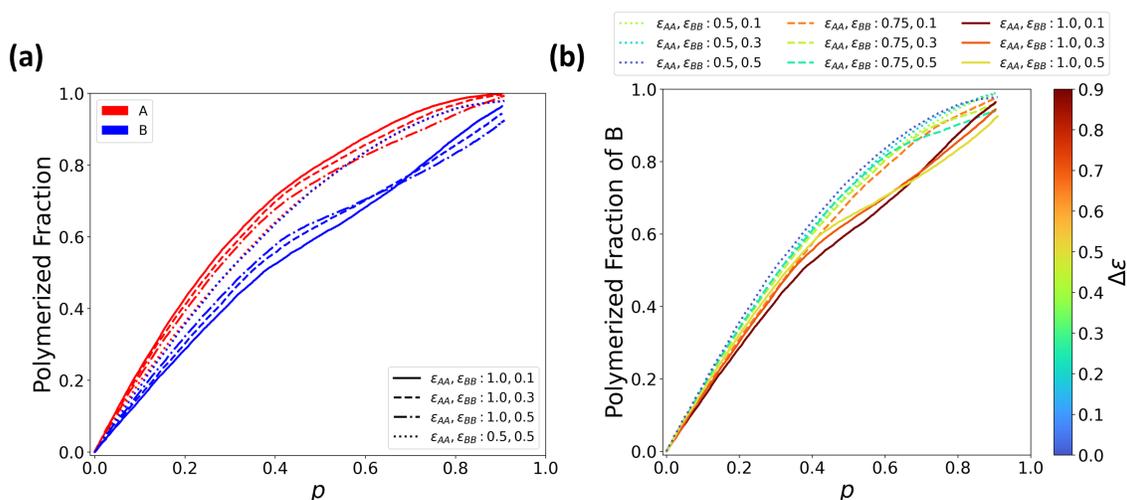
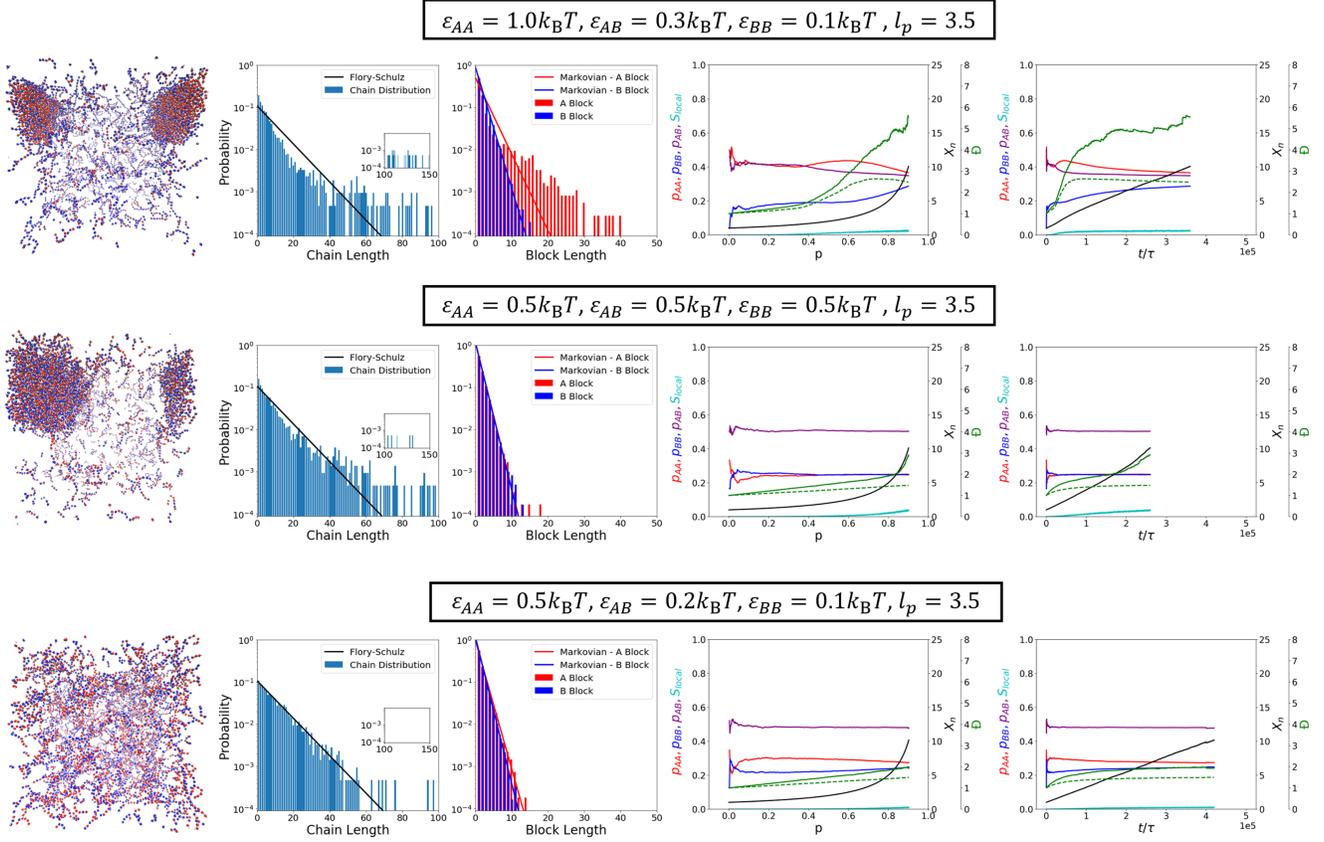


Figure B.4: **Polymerized fraction of monomer species.** (a) The fraction of **A** and **B** monomers which have polymerized as a function of the reaction extent,  $p$ , for stiff chains of  $l_p = 16.5$  and select attraction values. Monomer species is indicated by line color, and line styles indicate the set of attractive interactions. (b) The fraction of **B** monomers which have polymerized as a function of the reaction extent,  $p$ , for stiff chains of  $l_p = 16.5$  at all attraction combinations. Color indicates the degree of attraction asymmetry,  $\Delta\epsilon$ , and line styles indicate  $\epsilon_{AA}$  value. All data was obtained from three independent simulation trials for each parameter set. In each case, similar results were observed for flexible chains.

## **B.6 Supplemental kinetics and length distributions**

Additional figures providing kinetic parameters, chain and block length distributions, and characteristic snapshots of the final system state for select  $\varepsilon$  and  $l_p$  parameter sets are shown in Fig. B.5 and Fig. B.6 below.



**Figure B.5: Length distributions and kinetic parameters for flexible-chain copolymerizations.** Columns from left to right: the final system snapshot, chain length distribution, block length distributions, and several kinetic parameters versus reaction extent and simulation time. The chain length and block length distributions are represented as histograms, with the predicted distributions from Flory-Schulz and Markovian statistics plotted as lines. Kinetic plots include the pairwise bonding probabilities ( $p_{AA}$ ,  $p_{AB}$ , and  $p_{BB}$ ),  $\overline{S_{local}}$ ,  $X_n$ , and  $D$ . For  $D$  a dashed line is calculated for blocks, while the solid line is calculated for the entire chain. Rows represent different attraction strengths, as indicated, with flexible chains of  $l_p = 3.5$ . All data is collected from three independent simulation trials at each parameter set.

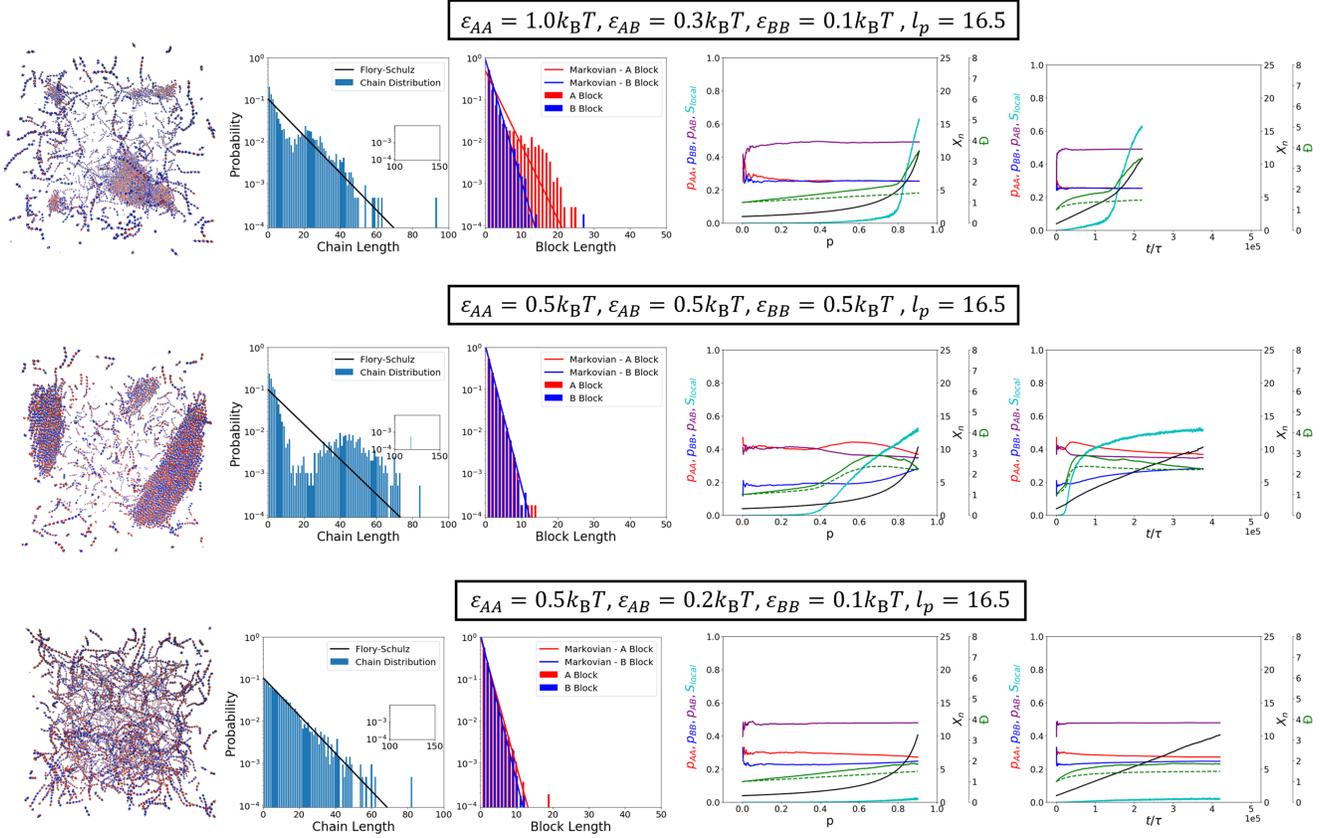


Figure B.6: **Length distributions and kinetic parameters for stiff-chain copolymerizations.** Columns from left to right: the final system snapshot, chain length distribution, block length distributions, and several kinetic parameters versus reaction extent and simulation time. The chain length and block length distributions are represented as histograms, with the predicted distributions from Flory-Schulz and Markovian statistics plotted as lines. Kinetic plots include the pairwise bonding probabilities ( $p_{AA}$ ,  $p_{AB}$ , and  $p_{BB}$ ),  $\overline{S}_{\text{local}}$ ,  $X_n$ , and  $D$ . For  $D$  a dashed line is calculated for blocks, while the solid line is calculated for the entire chain. Rows represent different attraction strengths, as indicated, with stiff chains of  $l_p = 16.5$ . All data is collected from three independent simulation trials at each parameter set.

# Appendix C

## Chapter 5 Supplemental Details and Results

### C.1 Simulation details

The model we use here was described and discussed in detail in Chapter 2. All simulations were performed in LAMMPS.<sup>228</sup> Additional details specific to Chapter 5 are described below.

#### C.1.1 Langevin dynamics and viscosity

Langevin dynamics was implemented in LAMMPS using the “fix langevin” command. A *damp* parameter is used to control the diffusion rate and the relaxation rate of the temperature, which is discussed in detail in the Supplemental Information of Ref. [59]. The values of *damp* and the associated viscosities explored in this work are shown in Table C.1. In the "standard" simulations in this work (marked with a \* in Table C.1, the *damp* parameter is set to  $0.1 \tau$ , resulting in a viscosity of  $\gamma = 0.1 \text{ mPa} \cdot \text{s}$ . For further discussion of this parameter, see Ref. [59].

#### C.1.2 Activation energies

To ensure that having a fixed value for  $\epsilon_{\text{rep}}$ , plus the additional soft repulsive potential between type 2 particles, was maintaining a consistent activation energy across monomer type pairs for our range of simulation parameters, we tabulated the activation energies for

$damp$ ( $\tau$ )	Viscosity (mPa·s)
0.01	1.00
0.03	0.3
0.05	0.2
0.1*	0.1*
0.3	0.03
0.5	0.02
1.0	0.01

Table C.1: **Simulation viscosity.** The first column shows the value of the Langevin parameter,  $damp$ , and the second the associated viscosity in units of mPa·s. The "standard" conditions are marked with a \*.

successful bond formations for a set of  $\sim 60000$  bonding events drawn from the full range of simulation parameters explored in this work. Activation energies are calculated from the total potential energy of the two bonding monomers arising from the LJ potential between type **1** particles, the soft repulsive potential between type **2** particles, and the intramonomer bond lengths and angles. The values of each of these contributions above their minimum energy are summed at the time of bond formation to obtain the total activation energy for the bonding event. Fig. C.1 shows the results of this comparison, broken down by monomer type pairs and attraction strength combination, demonstrating consistent activation energies across monomer type pairing and simulation parameters.

## C.2 Chain length dependence of bonding statistics

The chain-length dependent bonding behavior and its sensitivity to chain stiffness, discussed in Fig. 5.4 in the main text of Chapter 5, is further demonstrated in Fig. C.2. Fig. C.2 shows the proportion of new bond formation involving oligomers, which we define to be chains of length  $\geq 3$ . Both the nematic ordering transition and the reduction in chain length dispersity are unique features of chains with sufficient stiffness. Flexible chains remain disordered throughout the reaction, and dispersity continues to increase monotonically. As such, while the fraction of oligomers participating in bond formation is similar between flexible and stiff chains (Fig. C.2a), bond formation in which both of the reacting chains are oligomers are significantly less likely for stiff chains after nematic alignment occurs (Fig. C.2b).

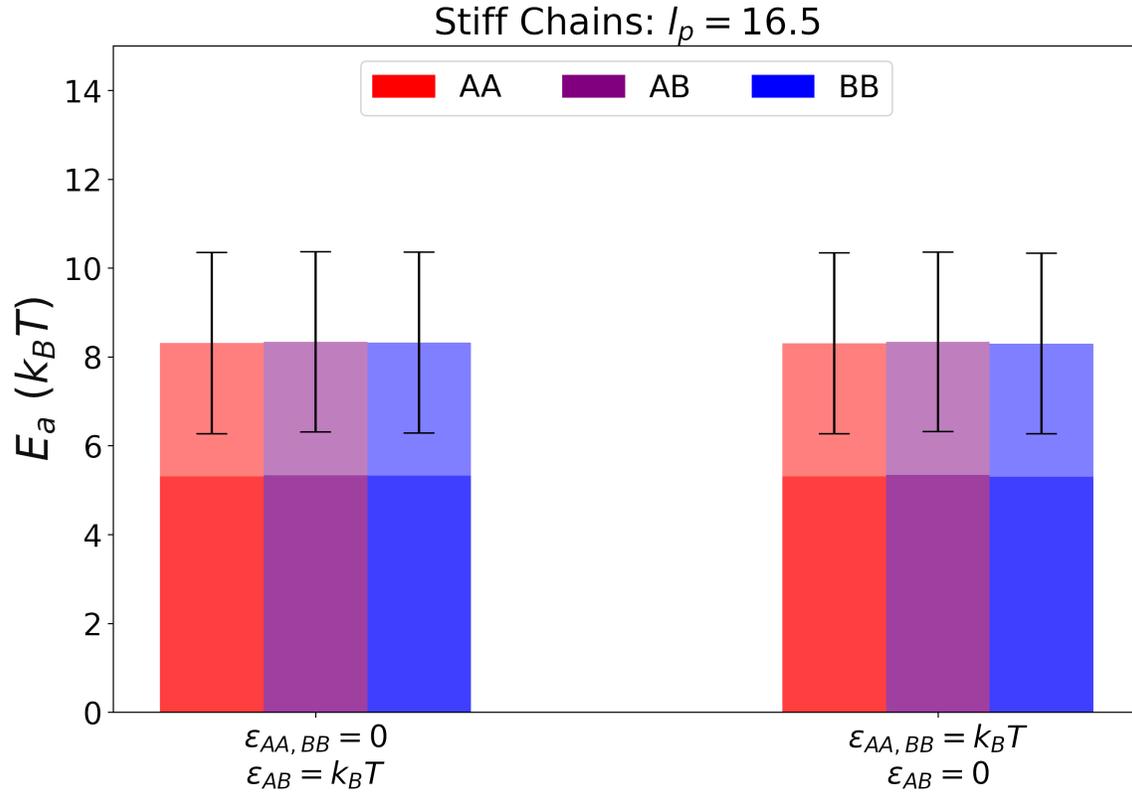
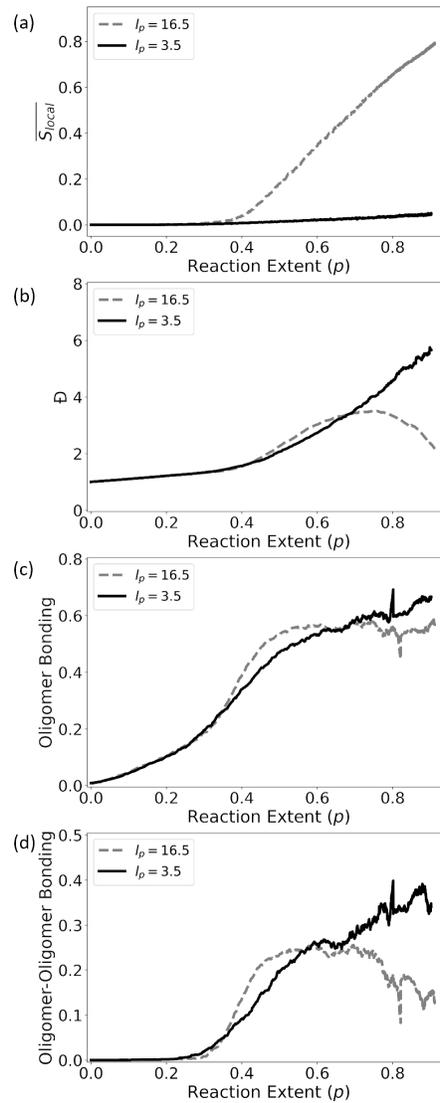


Figure C.1: **Activation energies.** The activation energies,  $E_a$ , encountered during the polymerization reaction. Bonding events were sampled from simulations under standard conditions and are shown here for two attraction strength combinations:  $\epsilon_{AA, BB} = k_B T$ ,  $\epsilon_{AB} = 0$  and  $\epsilon_{AA, BB} = 0$ ,  $\epsilon_{AB} = k_B T$ ). Sample populations are further separated by monomer pair, namely **AA**, **AB**, and **BB**. In total,  $\sim 60000$  bonding events were sampled. Mean values of the measured activation energies for each combination are shown, with the error bars representing the standard deviation. The dark shades of colors represent the constant component of activation energy, composed of LJ repulsions and geometric constraints, which has an average value of  $\sim 5.3k_B T$ . The lighter shaded region represents the short-ranged repulsion between type-2 particles from  $E_{\text{barr}}$ , which is varied across simulations to control the total activation energy.



**Figure C.2: Proportion of oligomers reacting.** (a) Local nematic ordering parameter,  $\overline{S}_{\text{local}}$ , as a function of reaction extent. (b) Chain length dispersity,  $\mathcal{D}$ , as a function of reaction extent. (c,d) The fractions of all bonding events occurring in which (c) one of the reacting chains is an oligomer and (d) both of the reacting chains are oligomers, are shown as a function of reaction extent  $\rho$ . Oligomers are defined as chains of three or more monomers. All results in (a-d) are shown for both flexible ( $l_p = 3.5$ ) and stiff ( $l_p = 16.5$ ) chains, and are obtained from three independent simulation trials for each chain stiffness.

### C.3 Deviations from Markovian predictions develop in time

To quantitatively explore the deviation from Markov statistics in our system, we compared the Markovian predictions for dispersity and block length distribution with the results of our simulations. We calculated the Wasserstein metric,<sup>96,97</sup>  $W_1$ , as a measure of the statistical distance between the Markovian distributions and our observed block length distributions.

### C.4 Clustering analysis

Identification of aggregates was performed via the use of the HDBSCAN<sup>122,229</sup> algorithm, which uses a density based criteria for cluster identification. At each time point chosen for analysis, HDBSCAN clustering was performed on the spatial coordinates of all monomer centers in the system, with a minimum cluster size of twelve monomers, chosen for consistency with our definition of  $\overline{S}_{\text{local}}$ . To ensure copolymer chains had consistent cluster definition, all monomers within a chain were assigned to a cluster in which any member of the chain was assigned. When any given chain had monomers belonging to more than one cluster, those clusters were merged and all members of each cluster were assigned to the new, merged cluster.

Additionally, the distribution of chain lengths and block lengths for both flexible and stiff chains were calculated for all chains belonging to a cluster, and all chains which remained unclustered. These distributions are shown in Fig. C.4&C.5 below, alongside the Wasserstein distance between the observed and predicted distributions.

To explore spatial differences in **A,B**-interface formation between flexible and stiff chains, we again employed a density based clustering approach, this time identifying the extent of clustering of bond types, namely **AA**, **AB**, and **BB** bonds. The proportions of spatially clustered bonds by bond type are shown in Fig. C.6, alongside the sequence neighbor probabilities,  $p_{AA}$ ,  $p_{AB}$ , and  $p_{BB}$ , *i.e.*, the proportion of **AA**, **AB**, and **BB** bonds respectively. Notably, in both flexible (Fig. C.6a) and stiff (Fig. C.6b) chains, the sequence neighbor probabilities,  $p_{ij}$ , behave nearly identically for  $p > 0.2$ , the range in which bond clustering begins and for which sufficient bonding has occurred for  $p_{ij}$  values to be well sampled. Likewise, the proportion of **AA** and **BB** bonds which are clustered behaves similarly for both flexible and stiff chains, demonstrating nearly complete spatial clustering as a consequence of the emergent phase-separation. In stiff chains, however, **AB** bonds show an earlier emergence of clustering, concomitant with the ordering transition, which reaches

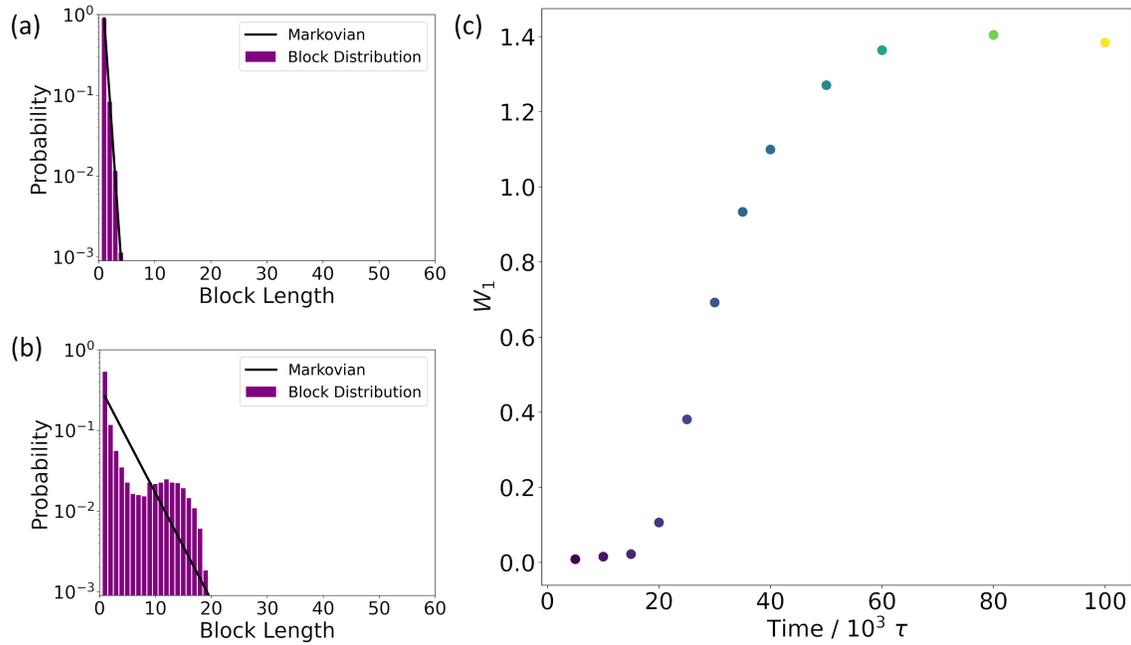


Figure C.3: **Block distribution deviation from Markovian statistics in time.** The observed block length distribution (discussed in Fig. 5.3b in the main text) at (a)  $5 \times 10^3 \tau$  and (b)  $100 \times 10^3 \tau$  are plotted alongside the distribution expected from Markov statistics. (c) The Wasserstein distance,  $W_1$ , is calculated between the observed block length distribution and the Markovian distribution and plotted as function of time, for points corresponding to the block distributions shown in Fig. 5.3b in the main text.

a significantly higher extent than that seen for flexible chains. This is a clear indication of the formation of densely co-located **AB** interfaces created when nematically ordered **A**-rich and **B**-rich aggregates merge.

## C.5 System parameters and characteristic timescale

Additional results obtained from variations to initial monomer density, non-bonded attraction strength, solvent viscosity, and activation energy are provided in Fig. C.8 below. In order to map these results to the characteristic timescales as discussed in Fig. 5.2 in the main text of Chapter 5, we first expressed these characteristic timescales in terms of the viscosity and activation energy parameters varied.

For the reactive timescale,  $\tau_R$ , we take the simplest definition<sup>121</sup> of  $\tau_R \equiv \frac{1}{k_{\text{eff}}}$ , where  $k_{\text{eff}}$  is the effective polymerization rate constant. In previous work,<sup>87</sup> we have worked out an expression for  $k_{\text{eff}}$  for our system in terms of the relative activation energies,  $E_A^{ij}$ , and the Arrhenius pre-exponential factors,  $A_{ij}$ , between each of the  $i, j$  monomer species pairs, namely:

$$k_{\text{eff}} = A_{AA, BB} \exp(-\beta E_a^{AA, BB}) + A_{AB} \exp(-\beta E_a^{AB}). \quad (\text{C.1})$$

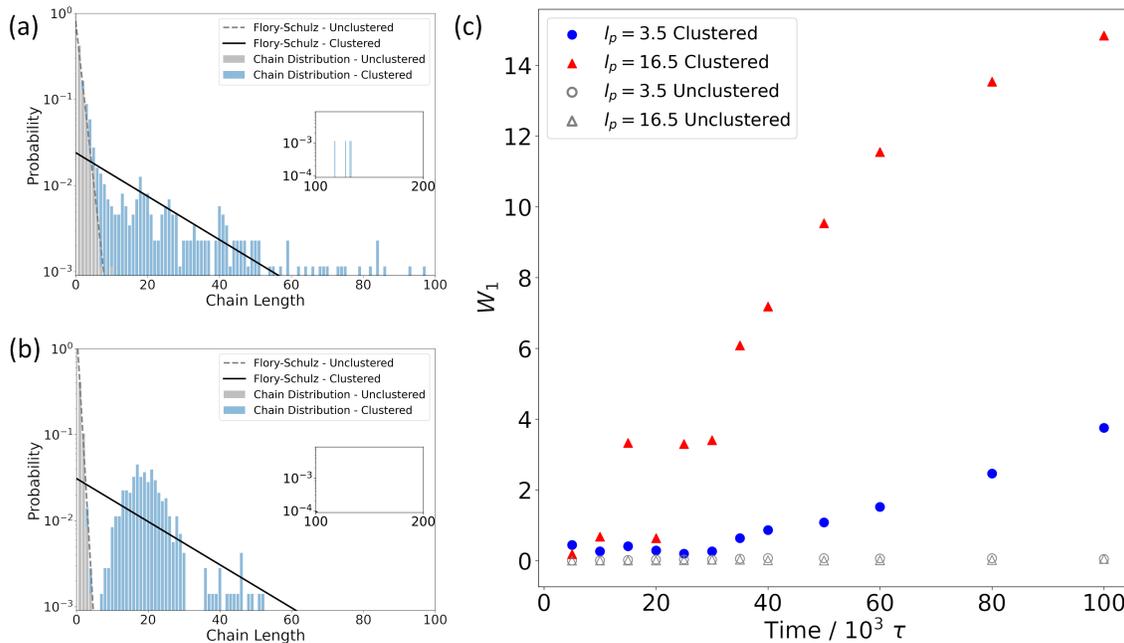
To define a diffusive timescale,  $\tau_D$ , we make use of the diffusion coefficient,  $D$ , obtained via the Stokes-Einstein relation along with the viscosity implicit in the Langevin equation we employ (discussed in Section C.1.1 above), taking

$$D = \frac{k_B T}{3\pi\eta d}. \quad (\text{C.2})$$

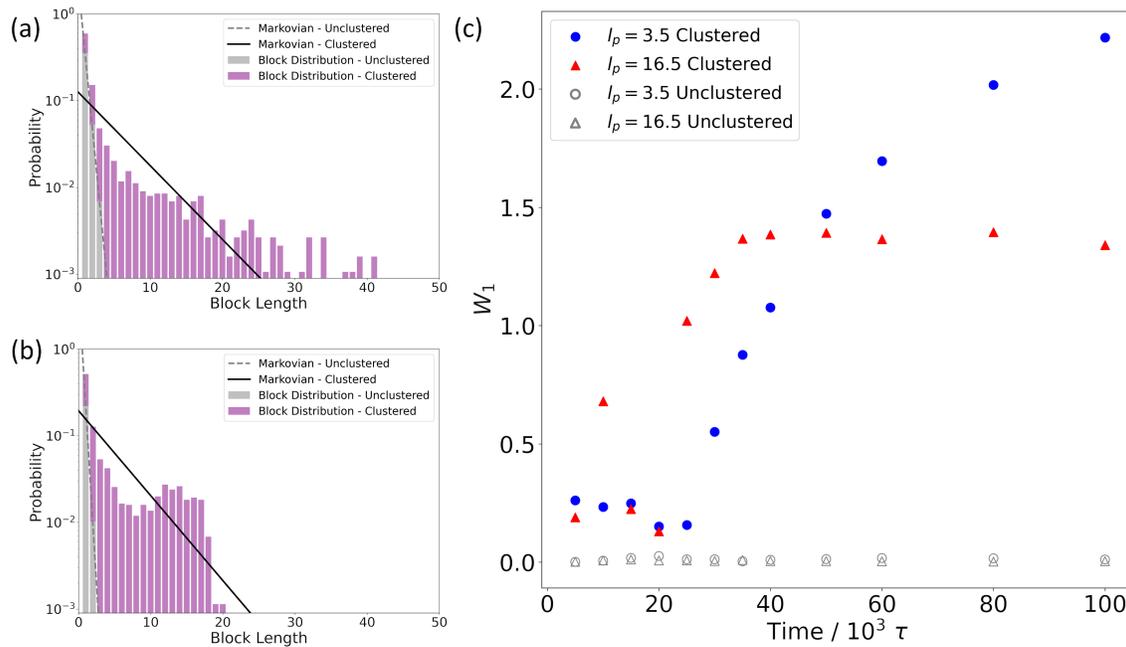
Here  $kT$  is the thermal energy,  $\eta$  is the solvent viscosity, and  $d$  is the diameter of a single monomer unit (namely  $d = 1\sigma$  in our reduced LJ unit system). From the diffusion coefficient, we define the characteristic diffusive timescale,  $\tau_D$ , in terms of the time for a monomer to diffuse its own diameter, that is:

$$\tau_D \equiv \frac{d^2}{D} = \frac{3\pi\eta d^3}{k_B T}. \quad (\text{C.3})$$

In addition to the response of block length distribution, nematic ordering, and dispersity to the timescales defined above, Fig. C.7 shows similar shifts in the extent and progression of oligomer and oligomer-oligomer bonding behaviors of systems in response to these



**Figure C.4: Chain distribution statistics for clustered and unclustered populations.** The observed chain length distribution at  $p = 0.9$  for (a) flexible chains ( $l_p = 3.5$ ) and (b) stiff chains ( $l_p = 16.5$ ) are plotted alongside the Flory-Schulz distribution. Chains were separated into clustered and unclustered populations based on the clustering analysis described above, with a corresponding reaction extent,  $p$ , within each population. The observed chain distribution is separated into clustered (blue) and unclustered (gray) portions, with the associated Flory-Schulz prediction obtained from the specific reaction extent of the population. (c) The Wasserstein distance,  $W_1$ , is calculated between the observed chain length distribution and the Flory-Schulz distribution and plotted as function of time, for points corresponding to the block distributions shown in Fig. 5.3b in the main text of Chapter 5. Results for both clustered and unclustered populations of both flexible ( $l_p = 3.5$ ) and stiff ( $l_p = 16.5$ ) chains are shown, as indicated by marker coloration and style.



**Figure C.5: Block distribution statistics for clustered and unclustered populations.** The observed block length distribution at  $p = 0.9$  for (a) flexible chains ( $l_p = 3.5$ ) and (b) stiff chains ( $l_p = 16.5$ ) are plotted alongside the distribution expected from Markov statistics. Sequences were separated into clustered and unclustered populations based on the clustering analysis described above, with a corresponding reaction extent,  $p$ , and value of  $p_{AA, BB}$  within each population. The observed block distribution is separated into clustered (purple) and unclustered (gray) portions, with the associated Markovian prediction obtained from the specific reaction extent and  $p_{AA, BB}$  of the population. (c) The Wasserstein distance,  $W_1$ , is calculated between the observed chain length distribution and the Markovian distribution and plotted as function of time, for points corresponding to the block distributions shown in Fig. 5.3b in the main text of Chapter 5. Results for both clustered and unclustered populations of both flexible ( $l_p = 3.5$ ) and stiff ( $l_p = 16.5$ ) chains are shown, as indicated by marker coloration and style.

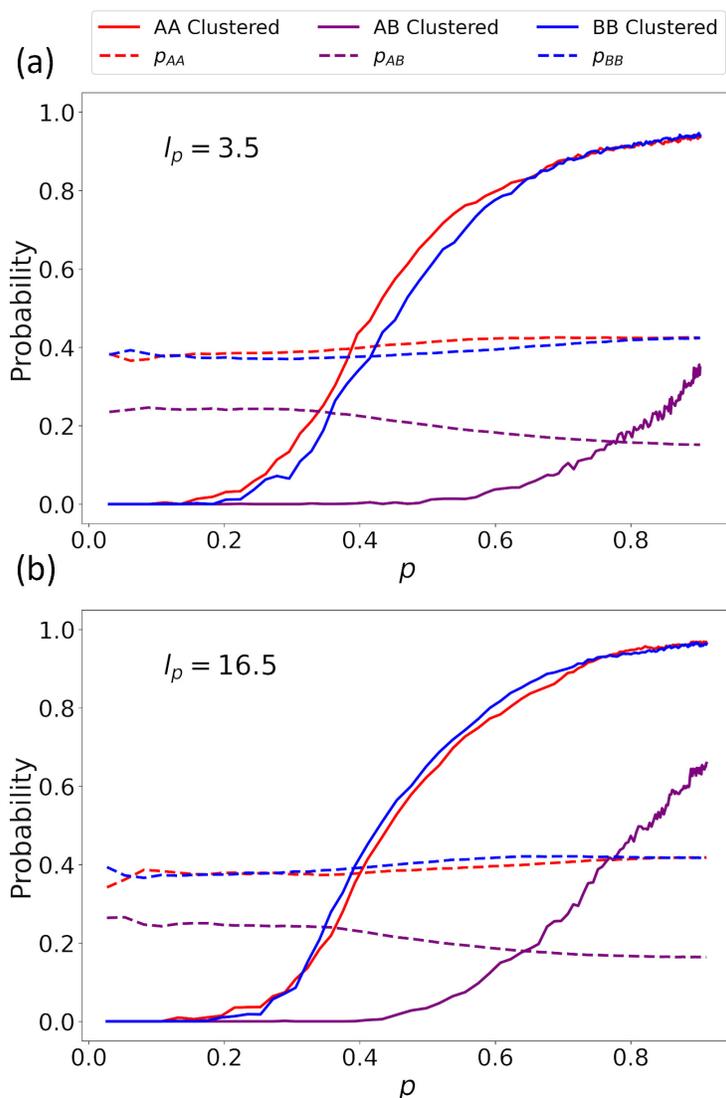


Figure C.6: **Spatial clustering of bonds by type.** Fractions of bonds formed that satisfy clustering criteria as a function of reaction extent for (a) flexible chains ( $l_p = 3.5$ ) and (b) stiff chains ( $l_p = 16.5$ ). Bonds were identified as clustered if there were at least 6 other bonds of the same type within  $2.5\sigma$ . Also shown are the sequence neighbor probabilities,  $p_{AA}$ ,  $p_{AB}$ , and  $p_{BB}$ : respectively, the probabilities of observing an **AA**, **AB**, or **BB** pair within all sequences in the system. Coloration indicates the bond type: red for **AA**, purple for **AB**, and blue for **BB**. Dotted lines are sequence neighbor probabilities and solid lines are the fraction of bonds clustered.

timescales.

## **C.6 Characteristic length in homopolymer system**

In order to verify that the characteristic length behavior we observed occurred irrespective of the phase-separation of comonomer species, we also ran homopolymer simulations under the same conditions as our standard copolymer simulation trials. The resulting chain length distributions, sampled from three independent simulation trials, are shown alongside the standard copolymer simulation in Fig. C.9.

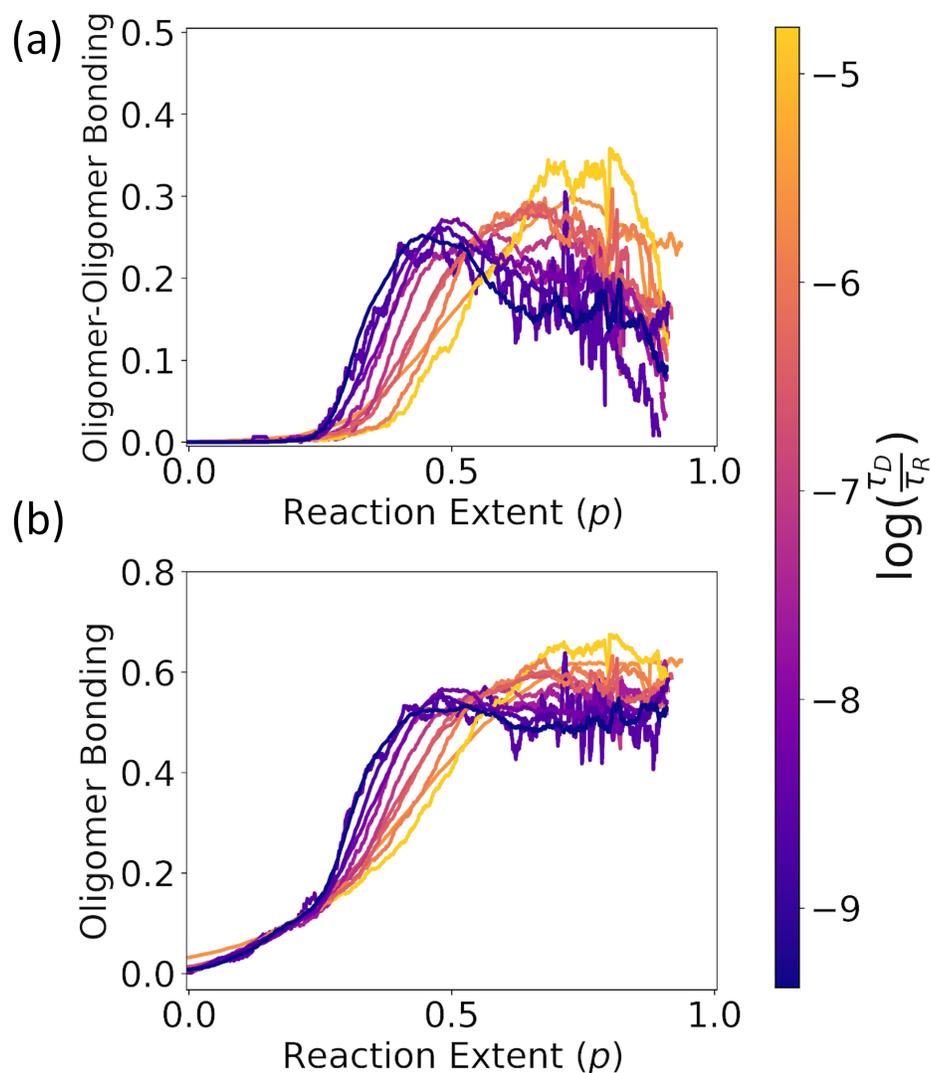
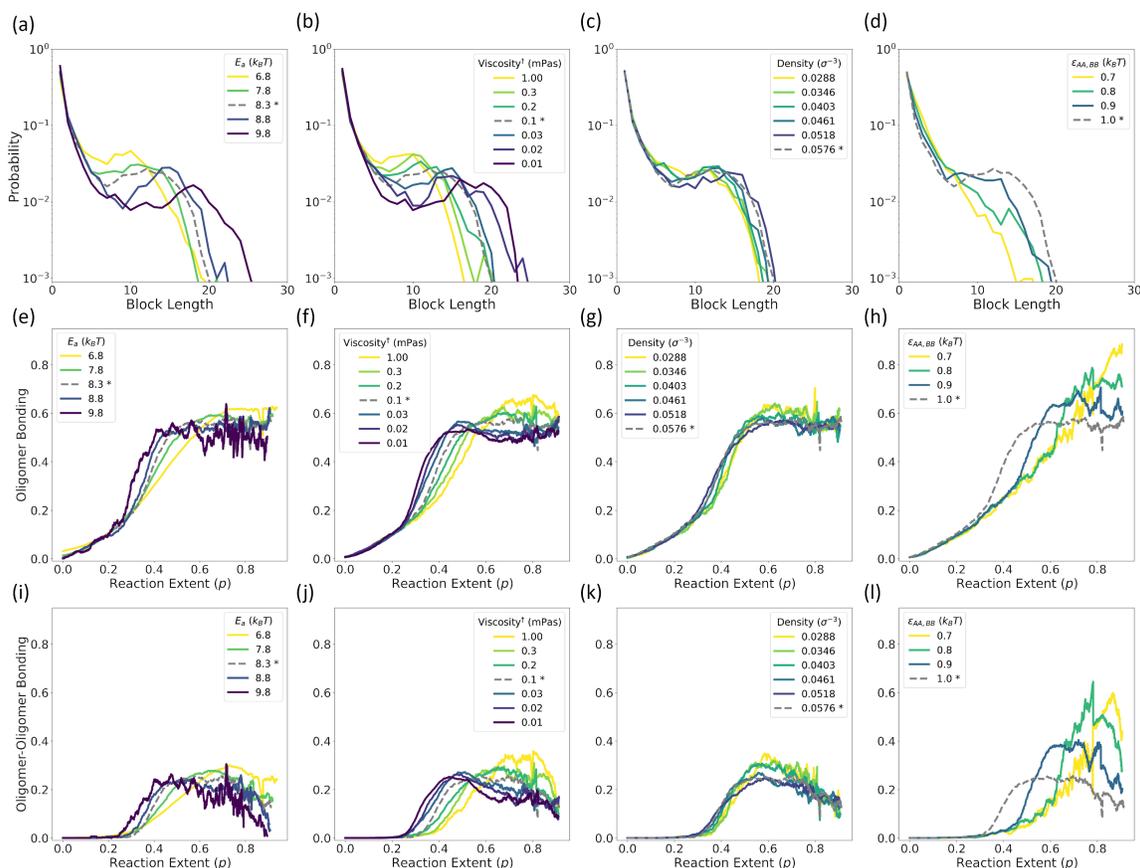


Figure C.7: **Oligomer bonding and characteristic timescales.** Oligomer bond formation for the simulations explored in Fig. 5.2 in the main text of Chapter 5. (a) Fraction of total new bond formation which is the result of oligomer bonding, i.e., of any chain of length  $\geq 3$  forming a bond, as a function of reaction extent. (b) Fraction of total new bond formation which is the result of oligomer-oligomer bonding, namely both reacting chains are of length  $\geq 3$  forming a bond, as a function of reaction extent. In both (a) and (b), coloration indicates the value of  $\log(\tau_D/\tau_R)$  the logarithm of the ratio of diffusive and reactive timescales for the simulation, with the dashed line indicating standard simulation conditions.



**Figure C.8: Block length distributions and bonding populations vary with activation energy, viscosity, monomer density, and attraction strength.** (a-d) Block length distributions of all-A or all-B blocks are shown for the copolymerization of stiff chains ( $l_p = 16.5$ ) at a reaction extent of  $p = 0.9$ . (e-h) Fraction of total new bond formation involving oligomers, as a function of reaction extent. Oligomers are defined here as chains of length three or longer. (i-l) Fraction of total new bond formation which is the result of oligomer-oligomer bonding, as a function of reaction extent. Results are shown here for variations in (a,e,i) activation energy, (b,f,j) solvent viscosity, (c,g,k) initial monomer densities, and (d,h,l) like-monomer attraction strengths. The gray, dashed lines in each plot, indicated with \* in the legends, are identical and display the block length distribution or oligomer-oligomer bonding fraction found under the standard conditions described in Ref. [59]. Bonding events are taken from three independent simulation trials per parameter set, and each distribution is obtained from three independent simulation trials at a reaction extent  $p = 0.9$ .

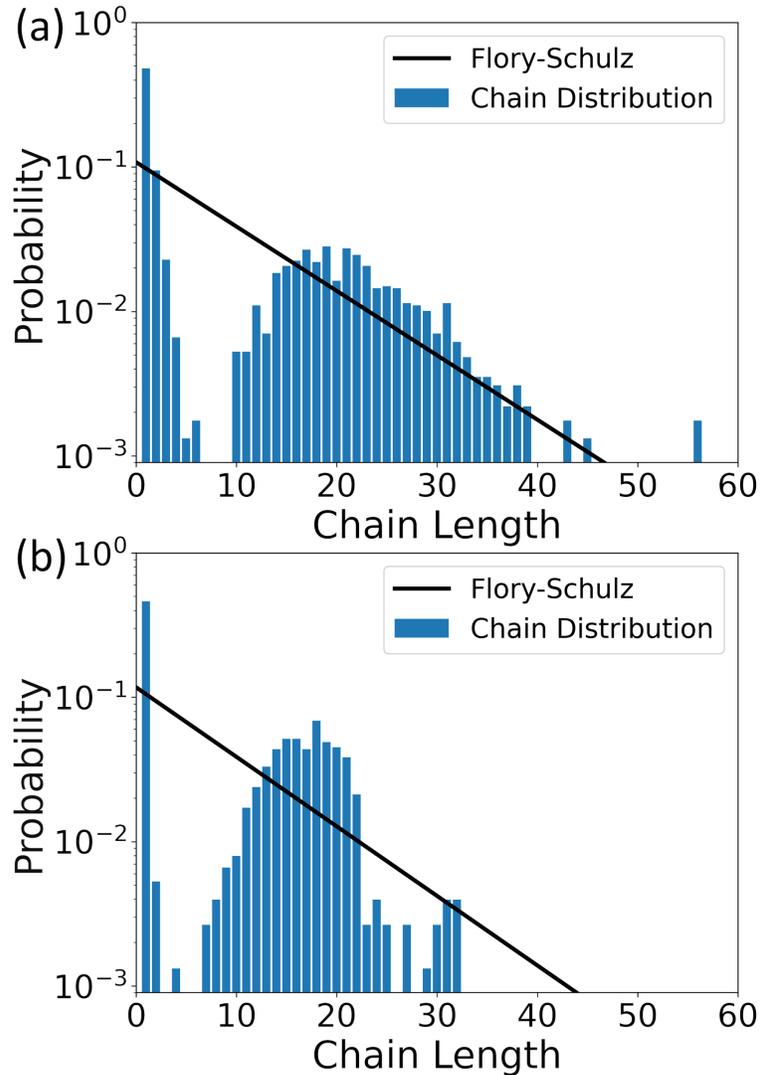


Figure C.9: **Chain length distribution shift persists throughout polymer composition changes.** The observed chain length distribution at  $p = 0.9$  for polymers with  $l_p = 16.5$  where  $\varepsilon_{AA, BB} = k_B T$  and  $\varepsilon_{AB} = 0$ , for (a) copolymer with equal fraction of monomer species **A** and **B**, and (b) a homopolymer of species **A** (equivalent to a homopolymer of **B** by symmetry). In both (a) and (b) the expected Flory-Schulz distribution is also shown.

# References

- (1) Staudinger, H. Über Polymerisation. *Berichte der deutschen chemischen Gesellschaft (A and B Series)* **1920**, *53*, 1073–1085.
- (2) Frey, H.; Johann, T. Celebrating 100 years of “polymer science”: Hermann Staudinger’s 1920 manifesto. *Polym. Chem.* **2020**, *11*, 8–14.
- (3) Abetz, V.; Chan, C. H.; Luscombe, C. K.; Matson, J. B.; Merna, J.; Nakano, T.; Raos, G.; Russell, G. T. Quo Vadis, Macromolecular Science? Reflections by the IUPAC Polymer Division on the Occasion of the Staudinger Centenary. *Israel Journal of Chemistry* **2020**, *60*, 9–19.
- (4) Feldman, D. Polymer History. *Designed Monomers and Polymers* **2008**, *11*, 1–15.
- (5) Young, R.; Lovell, P., *Introduction to Polymers: Third Edition*, 3rd ed.; CRC Press: United States, 2011.
- (6) Yang, G. G.; Choi, H. J.; Han, K. H.; Kim, J. H.; Lee, C. W.; Jung, E. I.; Jin, H. M.; Kim, S. O. Block Copolymer Nanopatterning for Nonsemiconductor Device Applications. *ACS Applied Materials & Interfaces* **2022**, *14*, PMID: 35230079, 12011–12037.
- (7) Pinto-Gómez, C.; Pérez-Murano, F.; Bausells, J.; Villanueva, L. G.; Fernández-Regúlez, M. Directed Self-Assembly of Block Copolymers for the Fabrication of Functional Devices. *Polymers* **2020**, *12*, DOI: 10.3390/polym12102432.
- (8) Sinturel, C.; Bates, F. S.; Hillmyer, M. A. High  $\chi$ -Low N Block Polymers: How Far Can We Go? *ACS Macro Letters* **2015**, *4*, PMID: 35596447, 1044–1050.
- (9) Lee, J. S.; Feijen, J. Polymersomes for drug delivery: Design, formation and characterization. *Journal of Controlled Release* **2012**, *161*, Drug Delivery Research in Europe, 473–483.

- (10) Fomina, N.; McFearin, C.; Sermsakdi, M.; Edigin, O.; Almutairi, A. UV and Near-IR Triggered Release from Polymeric Nanoparticles. *Journal of the American Chemical Society* **2010**, *132*, PMID: 20568765, 9540–9542.
- (11) Li, S.; Byrne, B.; Welsh, J.; Palmer, A. F. Self-assembled poly(butadiene)-b-poly(ethylene oxide) polymersomes as paclitaxel carriers. *Biotechnol Prog* **2007**, *23*, 278–285.
- (12) Discher, D. E.; Eisenberg, A. Polymer Vesicles. *Science* **2002**, *297*, 967 LP –973.
- (13) Harwood, H. J.; Ritchey, W. M. The characterization of sequence distribution in copolymers. *Journal of Polymer Science Part B: Polymer Letters* **1964**, *2*, 601–607.
- (14) Lutz, J.-F. Defining the Field of Sequence-Controlled Polymers. *Macromol. Rapid Commun.* **2017**, *38*, 1700582.
- (15) Berg, J. M.; Tymoczko, J. L.; J, G. J. G.; Stryer, L., *Biochemistry*, 8th; W. H. Freeman: New York, 2016.
- (16) Lutz, J.-F.; Ouchi, M.; Liu, D. R.; Sawamoto, M. Sequence-Controlled Polymers. *Science* **2013**, *341*, 1238149.
- (17) Ouchi, M.; Sawamoto, M. Sequence-Controlled Polymers via Reversible- Deactivation Radical Polymerization. *Polym. J.* **2017**, *50*, 83–94.
- (18) Xia, L.; Zhang, Z. Synthesis of Sequence-Controlled Polymers via Sequential Multicomponent Reactions and Interconvertible Hybrid Copolymerizations. *Polym. J.* **2020**, 33–43.
- (19) De Neve, J.; Haven, J. J.; Maes, L.; Junkers, T. Sequence-Definition from Controlled Polymerization: The Next Generation of Materials. *Polym. Chem.* **2018**, *9*, 4692–4705.
- (20) Amir, F.; Jia, Z.; Monteiro, M. J. Sequence Control of Macromers via Iterative Sequential and Exponential Growth. *J. Am. Chem. Soc.* **2016**, *138*, 16600–16603.
- (21) Wang, S.; Tao, Y.; Wang, J.; Tao, Y. A Versatile Strategy for the Synthesis of Sequence-Defined Peptoids with Side-Chain and Backbone Diversity via Amino Acid Building Blocks. *Chem. Sci.* **2019**, *10*, 1531–1538.
- (22) Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.

- (23) Chiefari, J.; Chong, Y. K. (; Ercole, F.; Krstina, J.; Jeffery, J.; Le, T. P. T.; Mayadunne, R. T. A.; Meijs, G. F.; Moad, C. L.; Moad, G.; Rizzardo, E.; Thang, S. H. Living Free-Radical Polymerization by Reversible Addition Fragmentation Chain Transfer: The RAFT Process. *Macromolecules* **1998**, *31*, 5559–5562.
- (24) Matyjaszewski, K.; Ziegler, M. J.; Arehart, S. V.; Greszta, D.; Pakula, T. Gradient Copolymers by Atom Transfer Radical Copolymerization. *J. Phys. Org. Chem.* **2000**, *13*, 775–786.
- (25) Moad, G.; Rizzardo, E.; Thang, S. H. Living Radical Polymerization by the RAFT Process. *Aust. J. Chem.* **2005**, *58*, 379–410.
- (26) Discekici, E. H.; Anastasaki, A.; Read De Alaniz, J.; Hawker, C. J. Evolution and Future Directions of Metal-Free Atom Transfer Radical Polymerization. *Macromolecules* **2018**, *51*, 7421–7434.
- (27) Truong, N. P.; Jones, G. R.; Bradford, K. G. E.; Konkolewicz, D.; Anastasaki, A. A Comparison of RAFT and ATRP Methods for Controlled Radical Polymerization. *Nat. Rev. Chem.* **2021**, *5*, 859–869.
- (28) Deacy, A. C.; Gregory, G. L.; Sulley, G. S.; Chen, T. T. D.; Williams, C. K. Sequence Control from Mixtures : Switchable Polymerization. *J. Am. Chem. Soc.* **2021**, *143*, 10021–10040.
- (29) DeStefano, A. J.; Segalman, R. A.; Davidson, E. C. Where Biology and Traditional Polymers Meet: The Potential of Associating Sequence-Defined Polymers for Materials Science. *JACS Au* **2021**, *1*, 1556–1571.
- (30) Yang, C.; Wu, K. B.; Deng, Y.; Yuan, J.; Niu, J. Geared Toward Applications: A Perspective on Functional Sequence-Controlled Polymers. *ACS Macro Lett.* **2021**, *10*, 243–257.
- (31) Rutten, M. G. T. A.; Vaandrager, F. W.; Elemans, J. A. A. W.; Nolte, R. J. M. Encoding Information into Polymers. *Nat. Rev. Chem.* **2018**, *2*, 365–381.
- (32) Lutz, J.-F. 100th Anniversary of Macromolecular Science Viewpoint: Toward Artificial Life-Supporting Macromolecules. *ACS Macro Lett.* **2020**, *9*, 185–189.
- (33) Church, G. M.; Gao, Y.; Kosuri, S. Next-Generation Digital Information Storage in DNA. *Science* **2012**, *337*, 1628.

- (34) Hoff, E. A.; De Hoe, G. X.; Mulvaney, C. M.; Hillmyer, M. A.; Alabi, C. A. Thiol–Ene Networks from Sequence-Defined Polyurethane Macromers. *J. Am. Chem. Soc.* **2020**, *142*, PMID: 32202773, 6729–6736.
- (35) McGraw, M. L.; Clarke, R. W.; Chen, E. Y.-X. Compounded Sequence Control in Polymerization of One-Pot Mixtures of Highly Reactive Acrylates by Differentiating Lewis Pairs. *J. Am. Chem. Soc.* **2020**, *142*, PMID: 32167755, 5969–5973.
- (36) Zhang, L.; Ji, Y.; Gu, X.; Zhang, W.; Zhou, N.; Zhang, Z.; Zhu, X. Synthesis of Sequence-Controlled Polymers with Pendent “Clickable” or Hydrophilic Groups via Latent Monomer Strategy. *React. Funct. Polym.* **2019**, *138*, 96–103.
- (37) Palermo, E. F.; McNeil, A. J. Impact of Copolymer Sequence on Solid-State Properties for Random, Gradient and Block Copolymers Containing Thiophene and Selenophene. *Macromolecules* **2012**, *45*, 5948–5955.
- (38) Aitken, B. S.; Buitrago, C. F.; Heffley, J. D.; Lee, M.; Gibson, H. W.; Winey, K. I.; Wagener, K. B. Precision Ionomers: Synthesis and Thermal/Mechanical Characterization. *Macromolecules* **2012**, *45*, 681–687.
- (39) Liu, D.; Lin, Y.; Bo, H.; Li, D.; Gong, K.; Zhang, Z.; Li, S. Effect of sequence distribution of block copolymers on the interfacial properties of ternary mixtures: a dissipative particle dynamics simulation. *RSC Adv.* **2022**, *12*, 3090–3096.
- (40) Ogura, Y.; Terashima, T.; Sawamoto, M. Synthesis of Fluorinated Gradient Copolymers via *in situ* Transesterification with Fluoroalcohols in Tandem Living Radical Polymerization. *Polym. Chem.* **2017**, *8*, 2299–2308.
- (41) Kim, J.; Mok, M. M.; Sandoval, R. W.; Woo, D. J.; Torkelson, J. M. Uniquely Broad Glass Transition Temperatures of Gradient Copolymers Relative to Random and Block Copolymers Containing Repulsive Comonomers. *Macromolecules* **2006**, *39*, 6152–6160.
- (42) Coldstream, J. G.; Camp, P. J.; Phillips, D. J.; Dowding, P. J. Gradient copolymers versus block copolymers: self-assembly in solution and surface adsorption. *Soft Matter* **2022**, *18*, 6538–6549.
- (43) Patterson, A. L.; Danielsen, S. P. O.; Yu, B.; Davidson, E. C.; Fredrickson, G. H.; Segalman, R. A. Sequence Effects on Block Copolymer Self-Assembly through Tuning Chain Conformation and Segregation Strength Utilizing Sequence-Defined Polypeptoids. *Macromolecules* **2019**, *52*, 1277–1286.

- (44) Golder, M. R.; Jiang, Y.; Teichen, P. E.; Nguyen, H. V.; Wang, W.; Milos, N.; Freedman, S. A.; Willard, A. P.; Johnson, J. A. Stereochemical Sequence Dictates Unimolecular Diblock Copolymer Assembly. *J. Am. Chem. Soc.* **2018**, *140*, 1596–1599.
- (45) Strickland, L. A.; Hall, C. K.; Genzer, J. Design of Copolymers with Tunable Randomness Using Discontinuous Molecular Dynamics Simulation. *Macromolecules* **2009**, *42*, 9063–9071.
- (46) Malik, R.; Hall, C. K.; Genzer, J. Effect of copolymer compatibilizer sequence on the dynamics of phase separation of immiscible binary homopolymer blends. *Soft Matter* **2011**, *7*, 10620–10630.
- (47) Rumyantsev, A. M.; Johner, A.; de Pablo, J. J. Sequence Blockiness Controls the Structure of Polyampholyte Necklaces. *ACS Macro Lett.* **2021**, *10*, 1048–1054.
- (48) Lytle, T. K.; Chang, L.-W.; Markiewicz, N.; Perry, S. L.; Sing, C. E. Designing Electrostatic Interactions via Polyelectrolyte Monomer Sequence. *ACS Cent. Sci.* **2019**, *5*, 709–718.
- (49) Chang, L.-W.; Lytle, T. K.; Radhakrishna, M.; Madinya, J. J.; Vélez, J.; Sing, C. E.; Perry, S. L. Sequence and entropy-based control of complex coacervates. *Nature Communications* **2017**, *8*, 1273.
- (50) Norris, B. N.; Zhang, S.; Campbell, C. M.; Auletta, J. T.; Calvo-Marzal, P.; Hutchison, G. R.; Meyer, T. Y. Sequence Matters: Modulating Electronic and Optical Properties of Conjugated Oligomers via Tailored Sequence. *Macromolecules* **2013**, *46*, 1384–1392.
- (51) Tsai, C.-H.; Fortney, A.; Qiu, Y.; Gil, R. R.; Yaron, D.; Kowalewski, T.; Noonan, K. J. T. Conjugated Polymers with Repeated Sequences of Group 16 Heterocycles Synthesized through Catalyst-Transfer Polycondensation. *J. Am. Chem. Soc.* **2016**, *138*, PMID: 27104362, 6798–6804.
- (52) Zhang, S.; Bauer, N. E.; Kanal, I. Y.; You, W.; Hutchison, G. R.; Meyer, T. Y. Sequence Effects in Donor–Acceptor Oligomeric Semiconductors Comprising Benzothiadiazole and Phenylenevinylene Monomers. *Macromolecules* **2017**, *50*, 151–161.

- (53) Mok, M. M.; Kim, J.; Torkelson, J. M. Gradient Copolymers with Broad Glass Transition Temperature Regions: Design of Purely Interphase Compositions for Damping Applications. *J. Polym. Sci., Part B: Polym. Phys.* **2008**, *46*, 48–58.
- (54) Luo, Y.; Guo, Y.; Gao, X.; Li, B. G.; Xie, T. A General Approach Towards Thermoplastic Multishape-Memory Polymers *via* Sequence Structure Design. *Adv. Mater.* **2013**, *25*, 743–748.
- (55) Mutlu, H.; Lutz, J.-F. Reading polymers: sequencing of natural and synthetic macromolecules. *Angew Chem Int Ed Engl* **2014**, *53*, 13010–13019.
- (56) Kuchanov, S. I.; Keshtov, M. L.; Halatur, P. G.; Vasnev, V. A.; Vinogradova, S. V.; Korshak, V. V. On the Principle of Equal Reactivity in Solution Polycondensation. *Die Makromolekulare Chemie* **1983**, *184*, 105–111.
- (57) Borisova, O. V.; Zaremski, M. Y.; Borisov, O. V.; Billon, L. The Well-Defined Bootstrap Effect in the Macroinitiator-Mediated Pseudoliving Radical Copolymerization of Styrene and Acrylic Acid. *Polymer Science - Series B* **2013**, *55*, 573–576.
- (58) Zhang, Z.; DuBay, K. H. Modeling the Influence of Emergent and Self-Limiting Phase Separations among Nascent Oligomers on Polymer Sequences Formed during Irreversible Step-Growth Copolymerizations. *Macromolecules* **2019**, *52*, 5480–5490.
- (59) Zhang, Z.; DuBay, K. H. The Sequence of a Step-Growth Copolymer Can Be Influenced by Its Own Persistence Length. *J. Phys. Chem. B* **2021**, *125*, 3426–3437.
- (60) Harwood, H. J. Structures and Compositions of Copolymers. *Makromol. Chem., Macromol. Symp.* **1987**, *10-11*, 331–354.
- (61) Coote, M. L.; Davis, T. P.; Klumperman, B.; Monteiro, M. J. A Mechanistic Perspective on Solvent Effects in Free-Radical Copolymerization. *J. Macromol. Sci., Part C: Polym. Rev.* **1998**, *38*, 567–593.
- (62) Kazantsev, O. A.; Kamorin, D. M.; Orekhov, D. V.; Sivokhin, A. P. Study of Amphiphilic Properties of Amine- and Oligo(Ethylene Glycol)-Containing (Meth)Acrylic Monomers. *Des. Monomers Polym.* **2015**, *18*, 378–384.
- (63) Gavrilov, A. A.; Chertovich, A. V. Copolymerization of Partly Incompatible Monomers: An Insight from Computer Simulations. *Macromolecules* **2017**, *50*, 4677–4685.

- (64) Kong, S.-M.; Liu, H.; Xue, Y.-H.; Liu, X.-L.; Jia, X.-X.; Cui, F.-C. Polymerization-induced polymer aggregation or polymer aggregation-enhanced polymerization? A computer simulation study. *Phys. Chem. Chem. Phys.* **2018**, *20*, 24379–24388.
- (65) Hao, J.; An, F.; Lu, C.; Liu, Y. Solvent Effects on Radical Copolymerization of Acrylonitrile and Methyl Acrylate: Solvent Polarity and Solvent-Monomer Interaction. *J. Macromol. Sci., Part A: Pure Appl. Chem.* **2019**, *56*, 1012–1021.
- (66) Penfold, N. J.; Yeow, J.; Boyer, C.; Armes, S. P. Emerging Trends in Polymerization-Induced Self-Assembly. *ACS Macro Lett.* **2019**, *8*, 1029–1054.
- (67) Canning, S. L.; Smith, G. N.; Armes, S. P. A Critical Appraisal of RAFT-Mediated Polymerization-Induced Self-Assembly. *Macromolecules* **2016**, *49*, 1985–2001.
- (68) Charleux, B.; Delaittre, G.; Rieger, J.; D'Agosto, F. Polymerization-Induced Self-Assembly: From Soluble Macromolecules to Block Copolymer Nano-Objects in One Step. *Macromolecules* **2012**, *45*, 6753–6765.
- (69) Chaduc, I.; Girod, M.; Antoine, R.; Charleux, B.; D'Agosto, F.; Lansalot, M. Batch Emulsion Polymerization Mediated by Poly(Methacrylic Acid) MacroRAFT Agents: One-Pot Synthesis of Self-Stabilized Particles. *Macromolecules* **2012**, *45*, 5881–5893.
- (70) Blanazs, A.; Madsen, J.; Battaglia, G.; Ryan, A. J.; Armes, S. P. Mechanistic Insights for Block Copolymer Morphologies: How Do Worms Form Vesicles? *J. Am. Chem. Soc.* **2011**, *133*, 16581–16587.
- (71) Zhang, W. J.; Hong, C. Y.; Pan, C. Y. Polymerization-Induced Self-Assembly of Functionalized Block Copolymer Nanoparticles and Their Application in Drug Delivery. *Macromol. Rapid Commun.* **2019**, *40*, 1–10.
- (72) Wang, X.; An, Z. New Insights into RAFT Dispersion Polymerization-Induced Self-Assembly: From Monomer Library, Morphological Control, and Stability to Driving Forces. *Macromol. Rapid Commun.* **2019**, *40*, 1–14.
- (73) Flory, P. J., *Principles of Polymer Chemistry*; Cornell University Press: Ithaca, New York, 1953; Chapter 3, pp 69–102.
- (74) Mayo, F. R.; Lewis, F. M. Copolymerization. I. A Basis for Comparing the Behavior of Monomers in Copolymerization; The Copolymerization of Styrene and Methyl Methacrylate. *J. Am. Chem. Soc.* **1944**, *66*, 1594–1601.

- (75) Carothers, W. H. Polymers and Polyfunctionality. *Trans. Faraday Soc.* **1936**, *32*, 39–49.
- (76) Doran, D.; Clarke, E.; Keenan, G.; Carrick, E.; Mathis, C.; Cronin, L. Exploring the sequence space of unknown oligomers and polymers. *Cell Reports Physical Science* **2021**, *2*, 100685.
- (77) Huang, F.; Lv, Y.; Wang, L.; Xu, P.; Lin, J.; Lin, S. An Insight into Polymerization-Induced Self-Assembly by Dissipative Particle Dynamics Simulation. *Soft Matter* **2016**, *12*, 6422–6429.
- (78) Xue, Y. H.; Zhu, Y. L.; Quan, W.; Qu, F. H.; Han, C.; Fan, J. T.; Liu, H. Polymer-grafted nanoparticles prepared by surface-initiated polymerization: The characterization of polymer chain conformation, grafting density and polydispersity correlated to the grafting surface curvature. *Physical Chemistry Chemical Physics* **2013**, *15*, 15356–15364.
- (79) Liu, H.; Li, M.; Lu, Z. Y.; Zhang, Z. G.; Sun, C. C. Influence of Surface-Initiated Polymerization Rate and Initiator Density on the Properties of Polymer Brushes. *Macromolecules* **2009**, *42*, 2863–2872.
- (80) Turgman-Cohen, S.; Genzer, J. Computer Simulation of Surface-Initiated Controlled Radical Polymerization: Effect of Free-Monomer Model on Brush Properties. *Macromol. Theory Simul.* **2019**, *28*, 1–10.
- (81) Farah, K.; Müller-Plathe, F.; Böhm, M. C. Classical Reactive Molecular Dynamics Implementations: State of the Art. *ChemPhysChem* **2012**, *13*, 1127–1151.
- (82) Farah, K.; Karimi-Varzaneh, H. A.; Müller-Plathe, F.; Böhm, M. C. Reactive Molecular Dynamics with Material-Specific Coarse-Grained Potentials: Growth of Polystyrene Chains from Styrene Monomers. *The Journal of Physical Chemistry B* **2010**, *114*, PMID: 20939547, 13656–13666.
- (83) Liu, H.; Zhu, Y.-L.; Lu, Z.-Y.; Müller-Plathe, F. A kinetic chain growth algorithm in coarse-grained simulations. *Journal of Computational Chemistry* **2016**, *37*, 2634–2646.
- (84) Krajniak, J.; Zhang, Z.; Pandiyan, S.; Nies, E.; Samaey, G. Coarse-grained molecular dynamics simulations of polymerization with forward and backward reactions. *Journal of Computational Chemistry* **2018**, *39*, 1764–1778.

- (85) Wang, L.; He, X. Reaction kinetics of nonideal hyperbranched polymerizations: Influences of chain rigidity and reaction reversibility. *Journal of Polymer Science Part A: Polymer Chemistry* **2012**, *50*, 2705–2714.
- (86) Hamblin, R. L.; Nguyen, N. Q.; DuBay, K. H. Selective Solvent Conditions Influence Sequence Development and Supramolecular Assembly in Step-Growth Copolymerization. *Soft Matter* **2022**, *18*, 943–955.
- (87) Nguyen, N. Q.; Hamblin, R. L.; DuBay, K. H. Emergent Sequence Biasing in Step-Growth Copolymerization: Influence of Non-Bonded Interactions and Comonomer Reactivities. *J. Phys. Chem. B* **2022**, *126*, 6585–6597.
- (88) Rouse, P. E. A Theory of the Linear Viscoelastic Properties of Dilute Solutions of Coiling Polymers. *J. Chem. Phys.* **1953**, *21*, 1272–1280.
- (89) Fritz, D.; Koschke, K.; Harmandaris, V. A.; van der Vegt, N. F. A.; Kremer, K. Multiscale modeling of soft matter: scaling of dynamics. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10412–10420.
- (90) Brown, W. B.; Gee, G. The Statistical Thermodynamics of Mixtures of Lennard-Jones Molecules I. Random Mixtures. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* **1957**, *250*, 175–220.
- (91) Kong, C. L. Combining Rules for Intermolecular Potential Parameters. II. Rules for the Lennard-Jones (12–6) Potential and the Morse Potential. *J. Chem. Phys.* **1973**, *59*, 2464–2467.
- (92) Lorentz, H. A. Ueber Die Anwendung Des Satzes Vom Virial in Der Kinetischen Theorie Der Gase. *Ann. Phys.* **1881**, *248*, 127–136.
- (93) Grønbech-Jensen, N. Complete Set of Stochastic Verlet-Type Thermostats for Correct Langevin Simulations. *Mol. Phys.* **2020**, *118*, e1662506.
- (94) Duering, E. R.; Kremer, K.; Grest, G. S. Structure and relaxation of end-linked polymer networks. *The Journal of Chemical Physics* **1994**, *101*, 8169–8192.
- (95) Arrhenius, S. On the Reaction Rate of the Inversion of Non-Refined Sugar Upon Souring. *Z. Phys. Chem.* **1941**, *4*, 226–248.
- (96) Dobrushin, R. L. Prescribing a System of Random Variables by Conditional Distributions. *Theory of Probability & Its Applications* **1970**, *15*, 458–486.

- (97) Vallender, S. S. Calculation of the Wasserstein Distance Between Probability Distributions on the Line. *Theory of Probability & Its Applications* **1974**, *18*, 784–786.
- (98) Smith, A. A.; Hall, A.; Wu, V.; Xu, T. Practical Prediction of Heteropolymer Composition and Drift. *ACS Macro Lett.* **2019**, *8*, 36–40.
- (99) Smith, A. A.; Autzen, H. E.; Laursen, T.; Wu, V.; Yen, M.; Hall, A.; Hansen, S. D.; Cheng, Y.; Xu, T. Controlling Styrene Maleic Acid Lipid Particles through RAFT. *Biomacromolecules* **2017**, *18*, 3706–3713.
- (100) Panganiban, B.; Qiao, B.; Jiang, T.; DelRe, C.; Obadia, M. M.; Nguyen, T. D.; Smith, A. A. A.; Hall, A.; Sit, I.; Crosby, M. G.; Dennis, P. B.; Drockenmuller, E.; de la Cruz, M. O.; Xu, T. Random Heteropolymers Preserve Protein Function in Foreign Environments. *Science* **2018**, *1243*, 1239–1243.
- (101) Semchivov, Y. D. Preferential Sorption of Monomers and Molecular Weight Effects in Radical Copolymerization. *Macromol. Symp.* **1996**, *111*, 317–328.
- (102) Warren, N. J.; Armes, S. P. Polymerization-Induced Self-Assembly of Block Copolymer Nano-Objects via RAFT Aqueous Dispersion Polymerization. *J. Am. Chem. Soc.* **2014**, *136*, 10174–10185.
- (103) Derry, M. J.; Fielding, L. A.; Armes, S. P. Industrially-Relevant Polymerization-Induced Self-Assembly Formulations in Non-Polar Solvents: RAFT Dispersion Polymerization of Benzyl Methacrylate. *Polym. Chem.* **2015**, *6*, 3054–3062.
- (104) Flory, P. J. Thermodynamics of High Polymer Solutions. *J. Chem. Phys.* **1941**, *10*, 51–61.
- (105) Huggins, M. L. Solutions of Long Chain Compounds. *J. Chem. Phys.* **1941**, *9*, 440.
- (106) Lin, H.-M.; Seaver, M.; Tang, K. Y.; Knight, A. E. W.; Parmenter, C. S. The Role of Intermolecular Potential Well Depths in Collision-Induced State Changes. *J. Chem. Phys.* **1979**, *70*, 5442–5457.
- (107) Chandler, D., *Introduction to Modern Statistical Mechanics*; Oxford University Press: New York, 1987.
- (108) Virtanen, P. et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261–272.
- (109) Pyun, C. W. Some Probability Relations in Multicomponent Copolymers. *J. Polym. Sci., Part A-2: Polym. Phys.* **1971**, *9*, 577–584.

- (110) Elias, H.-G. In *Macromolecules: Volume 2: Synthesis, Materials, and Technology*; Springer US: Boston, MA, 1984, pp 755–799.
- (111) Peng, C.-K.; Buldyrev, S. V.; Goldberger, A. L.; Havlin, S.; Sciortino, F.; Simons, M.; Stanley, H. Long-range correlations in nucleotide sequences. *Nature* **1992**, *356*, 168–170.
- (112) Flory, P. J. Molecular Size Distribution in Linear Condensation Polymers1. *J. Am. Chem. Soc.* **1936**, *58*, 1877–1885.
- (113) Heinz, D.; Amado, E.; Kressler, J. Polyphilicity—An Extension of the Concept of Amphiphilicity in Polymers. *Polymers* **2018**, *10*, 960.
- (114) Matsidik, R.; Komber, H.; Sommer, M. Rational Use of Aromatic Solvents for Direct Arylation Polycondensation: C–H Reactivity Versus Solvent Quality. *ACS Macro Lett.* **2015**, *4*, 1346–1350.
- (115) Plochocka, K. Effect of the Reaction Medium on Radical Copolymerization. *Journal of Macromolecular Science, Part C* **1981**, *20*, 67–148.
- (116) Flory, P. J. Statistical thermodynamics of semi-flexible chain molecules. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1956**, *234*, 60–73.
- (117) DiMarzio, E.; Gibbs, J. Chain stiffness and the lattice theory of polymer phases. *J. Chem. Phys.* **1958**, *28*, 807–813.
- (118) Fredrickson, G. H.; Liu, A. J.; Bates, F. S. Entropic Corrections to the Flory-Huggins Theory of Polymer Blends: Architectural and Conformational Effects. *Macromolecules* **1994**, *27*, 2503–2511.
- (119) Zierenberg, J.; Janke, W. From Amorphous Aggregates to Polymer Bundles: The Role of Stiffness on Structural Phases in Polymer Aggregation. *EPL* **2015**, *109*, 28002.
- (120) Midya, J.; Egorov, S. A.; Binder, K.; Nikoubashman, A. Phase Behavior of Flexible and Semiflexible Polymers in Solvents of Varying Quality. *J. Chem. Phys.* **2019**, *151*, 034902.
- (121) Wartha, E.-M.; Bösenhofer, M.; Harasek, M. Characteristic Chemical Time Scales for Reactive Flow Modeling. *Combust. Sci. Technol.* **2021**, *193*, 2807–2832.
- (122) McInnes, L.; Healy, J.; Astels, S. HDBSCAN: Hierarchical Density Based Clustering. *JOSS* **2017**, *2*, DOI: 10.21105/joss.00205.

- (123) Mabesoone, M. F. J.; Palmans, A. R. A.; Meijer, E. W. Solute–Solvent Interactions in Modern Physical Organic Chemistry: Supramolecular Polymers as a Muse. *J. Am. Chem. Soc.* **2020**, *142*, PMID: 33174741, 19781–19798.
- (124) Cruz, V. L.; Muñoz–Escalona, A.; Martínez–Salazar, J. A theoretical study of the comonomer effect in the ethylene polymerization with zirconocene catalytic systems. *Journal of Polymer Science Part A: Polymer Chemistry* **1998**, *36*, 1157–1167.
- (125) Satoh, K.; Ozawa, S.; Mizutani, M.; Nagai, K.; Kamigaito, M. Sequence-regulated vinyl copolymers by metal-catalysed step-growth radical polymerization. *Nature Communications* **2010**, *1*, 6.
- (126) Ramli, R. A.; Laftah, W. A.; Hashim, S. Core–shell polymers: a review. *RSC Adv.* **2013**, *3*, 15543–15565.
- (127) Lovell, P. A.; Schork, F. J. Fundamentals of Emulsion Polymerization. *Biomacromolecules* **2020**, *21*, PMID: 32543173, 4396–4441.
- (128) Kravchenko, V. S.; Potemkin, I. I. Micelles of Gradient Vs Diblock Copolymers: Difference in the Internal Structure and Properties. *J. Phys. Chem. B* **2016**, *120*, PMID: 27933941, 12211–12217.
- (129) Kravchenko, V. S.; Abetz, V.; Potemkin, I. I. Self-Assembly of Gradient Copolymers in a Selective Solvent. New Structures and Comparison with Diblock and Statistical Copolymers. *Polymer* **2021**, *235*, 124288.
- (130) Pandav, G.; Pryamitsyn, V.; Gallow, K. C.; Loo, Y.-L.; Genzer, J.; Ganesan, V. Phase behavior of gradient copolymer solutions: a Monte Carlo simulation study. *Soft Matter* **2012**, *8*, 6471–6482.
- (131) Gross, Y. M.; Ludwigs, S. P(NDI2OD-T2) Revisited – Aggregation Control as Key for High Performance N-Type Applications. *Synth. Met.* **2019**, *253*, 73–87.
- (132) Kos, P. I.; Ivanov, V. A.; Chertovich, A. V. Crystallization of semiflexible polymers in melts and solutions. *Soft Matter* **2021**, *17*, 2392–2403.
- (133) Polovnikov, K. E.; Potemkin, I. I. Effect of Architecture on Micelle Formation and Liquid-Crystalline Ordering in Solutions of Block Copolymers Comprising Flexible and Rigid Blocks: Rod–Coil Vs Y-Shaped Vs Comblike Copolymers. *J. Phys. Chem. B* **2017**, *121*, PMID: 28985085, 10180–10189.
- (134) Wang, L.; Ding, Y.; Liu, Q.; Zhao, Q.; Dai, X.; Lu, X.; Cai, Y. Sequence-Controlled Polymerization-Induced Self-Assembly. *ACS Macro Lett.* **2019**, *8*, 623–628.

- (135) Xu, S.; Zhang, T.; Kuchel, R. P.; Yeow, J.; Boyer, C. Gradient Polymerization-Induced Self-Assembly: A One-Step Approach. *Macromol. Rapid Commun.* **2020**, *41*, 1900493.
- (136) Lesage de la Haye, J.; Zhang, X.; Chaduc, I.; Brunel, F.; Lansalot, M.; D'Agosto, F. The Effect of Hydrophile Topology in RAFT-Mediated Polymerization-Induced Self-Assembly. *Angew. Chem. Int. Ed.* **2016**, *55*, 3739–3743.
- (137) Brunel, F.; Lesage De La Haye, J.; Lansalot, M.; D'Agosto, F. New Insight into Cluster Aggregation Mechanism during Polymerization-Induced Self-Assembly by Molecular Dynamics Simulation. *J. Phys. Chem. B* **2019**, *123*, 6609–6617.
- (138) Masser, K. A.; Bain, E. D.; Beyer, F. L.; Savage, A. M.; Yu, J. H.; Lenhart, J. L. Influence of Nano-Scale Morphology on Impact Toughness of Epoxy Blends. *Polymer* **2016**, *103*, New Polymeric Materials and Characterization Methods for Water Purification, 337–346.
- (139) Stevens, M. J. Simulation of Polymerization Induced Phase Separation in Model Thermosets. *J. Chem. Phys.* **2021**, *155*, 1–9.
- (140) Biedermann, F.; Schneider, H.-J. Experimental Binding Energies in Supramolecular Complexes. *Chem. Rev.* **2016**, *116*, PMID: 27136957, 5216–5300.
- (141) Driver, M. D.; Williamson, M. J.; Cook, J. L.; Hunter, C. A. Functional group interaction profiles: a general treatment of solvent effects on non-covalent interactions. *Chem. Sci.* **2020**, *11*, 4456–4466.
- (142) Hunter, C. A. Quantifying Intermolecular Interactions: Guidelines for the Molecular Recognition Toolbox. *Angew. Chem. Int. Ed.* **2004**, *43*, 5310–5324.
- (143) Ramachandran, R.; Beaucage, G.; Kulkarni, A. S.; McFaddin, D.; Merrick-Mack, J.; Galiatsatos, V. Persistence length of short-chain branched polyethylene. *Macromolecules* **2008**, *41*, 9802–9806.
- (144) Gettinger, C. L.; Heeger, A. J.; Drake, J. M.; Pine, D. J. A Photoluminescence Study of Poly(Phenylene Vinylene) Derivatives: The Effect of Intrinsic Persistence Length. *J. Chem. Phys.* **1994**, *101*, 1673–1678.
- (145) Brinkers, S.; Dietrich, H. R.; de Groote, F. H.; Young, I. T.; Rieger, B. The persistence length of double stranded DNA determined using dark field tethered particle motion. *The Journal of chemical physics* **2009**, *130*, 06B607.

- (146) Walters, R. H.; Murphy, R. M. Examining Polyglutamine Peptide Length: A Connection between Collapsed Conformations and Increased Aggregation. *J. Mol. Biol.* **2009**, *393*, 978–992.
- (147) Chin, A. F.; Toptygin, D.; Elam, W. A.; Schrank, T. P.; Hilser, V. J. Phosphorylation Increases Persistence Length and End-To-End Distance of a Segment of Tau Protein. *Biophys. J.* **2016**, *110*, 362–371.
- (148) Debie, C.; Coudert, N.; Abdul, J.; Harrisson, S.; Colombani, O.; Rieger, J. Controlling the Composition Profile of Acrylic Acid Copolymers by Tuning the pH of Polymerization in Aqueous Dispersed Media. *Macromolecules* **2023**, *56*, 8497–8506.
- (149) Kashchiev, D.; Vekilov, P. G.; Kolomeisky, A. B. Kinetics of Two-Step Nucleation of Crystals. *J. Chem. Phys.* **2005**, *122*, 244706.
- (150) Erdemir, D.; Lee, A. Y.; Myerson, A. S. Nucleation of Crystals from Solution: Classical and Two-Step Models. *Acc. Chem. Res.* **2009**, *42*, PMID: 19402623, 621–629.
- (151) Yoreo, J. J. D.; Gilbert, P. U. P. A.; Sommerdijk, N. A. J. M.; Penn, R. L.; Whitelam, S.; Joester, D.; Zhang, H.; Rimer, J. D.; Navrotsky, A.; Banfield, J. F.; Wallace, A. F.; Michel, F. M.; Meldrum, F. C.; Cölfen, H.; Dove, P. M. Crystallization by Particle Attachment in Synthetic, Biogenic, and Geologic Environments. *Science* **2015**, *349*, aaa6760.
- (152) Zryd, J. L.; Burghardt, W. R. Phase Separation, Crystallization, and Structure Formation in Immiscible Polymer Solutions. *J. Appl. Polym. Sci.* **1995**, *57*, 1525–1537.
- (153) Galkin, O.; Vekilov, P. G. Control of Protein Crystal Nucleation around the Metastable Liquid-Liquid Phase Boundary. *Proceedings of the National Academy of Sciences* **2000**, *97*, 6277–6281.
- (154) Hu, W.; Frenkel, D.; Mathot, V. B. F. Simulation of Shish-Kebab Crystallite Induced by a Single Prealigned Macromolecule. *Macromolecules* **2002**, *35*, 7172–7174.
- (155) Hu, W.; Frenkel, D. In *Interphases and Mesophases in Polymer Crystallization III*, Allegra, G., Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2005, pp 1–35.
- (156) Haxton, T. K.; Whitelam, S. Design rules for the self-assembly of a protein crystal. *Soft Matter* **2012**, *8*, 3558–3562.

- (157) Haxton, T. K.; Hedges, L. O.; Whitlam, S. Crystallization and arrest mechanisms of model colloids. *Soft Matter* **2015**, *11*, 9307–9320.
- (158) Hill, T. L. A Theoretical Study of Cooperative Dual Linear Aggregation and the Vernier Effect. *Biophys. Chem.* **1986**, *25*, 1–15.
- (159) Kumar, R.; Liu, Z.; Lokitz, B.; Chen, J.; Carrillo, J.-M.; Jakowski, J.; Collier, C. P.; Retterer, S.; Advincula, R. Harnessing autocatalytic reactions in polymerization and depolymerization. *MRS Commun.* **2021**, *11*, 377–390.
- (160) Norrish, R. G. W.; Smith, R. R. Catalysed Polymerization of Methyl Methacrylate in the Liquid Phase. *Nature* **1942**, *150*, 336–337.
- (161) O’Shaughnessy, B.; Yu, J. Autoacceleration in Free Radical Polymerization. *Phys. Rev. Lett.* **1994**, *73*, 1723–1726.
- (162) Peng, J.; Han, Y. Recent advances in conjugated polythiophene-based rod–rod block copolymers: From morphology control to optoelectronic applications. *Giant* **2020**, *4*, 100039.
- (163) Scherf, U.; Gutacker, A.; Koenen, N. All-Conjugated Block Copolymers. *Accounts of Chemical Research* **2008**, *41*, PMID: 18402471, 1086–1097.
- (164) Danielsen, S. P. O.; Bridges, C. R.; Segalman, R. A. Chain Stiffness of Donor–Acceptor Conjugated Polymers in Solution. *Macromolecules* **2022**, *55*, 437–449.
- (165) Bronstein, H.; Nielsen, C. B.; Schroeder, B. C.; McCulloch, I. The role of chemical design in the performance of organic semiconductors. *Nature Reviews Chemistry* **2020**, *4*, 66–77.
- (166) Danielsen, S. P. O.; Davidson, E. C.; Fredrickson, G. H.; Segalman, R. A. Absence of Electrostatic Rigidity in Conjugated Polyelectrolytes with Pendant Charges. *ACS Macro Letters* **2019**, *8*, PMID: 35619444, 1147–1152.
- (167) Guo, X.; Baumgarten, M.; Müllen, K. Designing  $\pi$ -conjugated polymers for organic electronics. *Progress in Polymer Science* **2013**, *38*, Topical issue on Conductive Polymers, 1832–1908.
- (168) Yu, H.; Li, S.; Schwieter, K. E.; Liu, Y.; Sun, B.; Moore, J. S.; Schroeder, C. M. Charge Transport in Sequence-Defined Conjugated Oligomers. *Journal of the American Chemical Society* **2020**, *142*, PMID: 32069403, 4852–4861.

- (169) Zhang, S.; Ma, L.; Ye, L.; Qin, Y.; Xu, Y.; Liu, X.; Wu, Y.; Zhao, W.; Ade, H.; Yao, H.; Hou, J. Modulation of Building Block Size in Conjugated Polymers with D–A Structure for Polymer Solar Cells. *Macromolecules* **2019**, *52*, 7929–7938.
- (170) Zhang, S.; Bauer, N. E.; Kanal, I. Y.; You, W.; Hutchison, G. R.; Meyer, T. Y. Sequence Effects in Donor–Acceptor Oligomeric Semiconductors Comprising Benzothiadiazole and Phenylenevinylene Monomers. *Macromolecules* **2017**, *50*, 151–161.
- (171) Alowakennu, M.; Omoniyi, A. O.; Ejeromedoghene, O.; Alli, Y. A.; Akor, E.; Nnyia, M. O. Harnessing photo-induced processes for the fabrication and application of functional conjugated and conducting polymer-based materials. *Journal of Molecular Structure* **2023**, *1292*, 136149.
- (172) Yilmaz, G.; Yagci, Y. Light-induced step-growth polymerization. *Progress in Polymer Science* **2020**, *100*, 101178.
- (173) Sakamoto, J.; Rehahn, M.; Wegner, G.; Schlüter, A. D. Suzuki Polycondensation: Polyarylenes à la Carte. *Macromolecular Rapid Communications* **2009**, *30*, 653–687.
- (174) Carsten, B.; He, F.; Son, H. J.; Xu, T.; Yu, L. Stille Polycondensation for Synthesis of Functional Materials. *Chemical Reviews* **2011**, *111*, PMID: 21314190, 1493–1528.
- (175) Kang, S.; Yoon, T. W.; Kim, G.-Y.; Kang, B. Review of Conjugated Polymer Nanoparticles: From Formulation to Applications. *ACS Applied Nano Materials* **2022**, *5*, 17436–17460.
- (176) Kim, C.; Joung, H.; Kim, H. J.; Paeng, K.; Kaufman, L. J.; Yang, J. Aggregates of conjugated polymers: bottom-up control of mesoscopic morphology and photo-physics. *NPG Asia Materials* **2023**, *15*, 32.
- (177) Noriega, R.; Rivnay, J.; Vandewal, K.; Koch, F. P. V.; Stingelin, N.; Smith, P.; Toney, M. F.; Salleo, A. A general relationship between disorder, aggregation and charge transport in conjugated polymers. *Nature Materials* **2013**, *12*, 1038–1044.
- (178) Ostrowski, D. P.; Lytwak, L. A.; Mejia, M. L.; Stevenson, K. J.; Holliday, B. J.; Vanden Bout, D. A. The Effects of Aggregation on Electronic and Optical Properties of Oligothiophene Particles. *ACS Nano* **2012**, *6*, PMID: 22590975, 5507–5513.

- (179) Guseva, E.; Zuckermann, R. N.; Dill, K. A. Foldamer hypothesis for the growth and sequence differentiation of prebiotic polymers. *Proceedings of the National Academy of Sciences* **2017**, *114*, E7460–E7468.
- (180) Wu, M.; Higgs, P. G. Origin of Self-Replicating Biopolymers: Autocatalytic Feedback Can Jump-Start the RNA World. *Journal of Molecular Evolution* **2009**, *69*, 541–554.
- (181) Lee, D. H.; Granja, J. R.; Martinez, J. A.; Severin, K.; Ghadiri, M. R. A self-replicating peptide. *Nature* **1996**, *382*, 525–528.
- (182) Rubinov, B.; Wagner, N.; Rapaport, H.; Ashkenasy, G. Self-Replicating Amphiphilic  $\beta$ -Sheet Peptides. *Angewandte Chemie International Edition* **2009**, *48*, 6683–6686.
- (183) Ohtsuki, H.; Nowak, M. A. Prolife catalysts and replicators. *Proceedings of the Royal Society B: Biological Sciences* **2009**, *276*, 3783–3790.
- (184) Chen, I. A.; Nowak, M. A. From Prolife to Life: How Chemical Kinetics Become Evolutionary Dynamics. *Accounts of Chemical Research* **2012**, *45*, PMID: 22335792, 2088–2096.
- (185) Derr, J.; Manapat, M. L.; Rajamani, S.; Leu, K.; Xulvi-Brunet, R.; Joseph, I.; Nowak, M. A.; Chen, I. A. Prebiotically plausible mechanisms increase compositional diversity of nucleic acid sequences. *Nucleic Acids Research* **2012**, *40*, 4711–4722.
- (186) Szostak, J.; Ellington, A. In Vitro Selection of Functional Nucleic Acids in the RNA World, 1993.
- (187) Lambert, J.-F. Adsorption and Polymerization of Amino Acids on Mineral Surfaces: A Review. *Origins of Life and Evolution of Biospheres* **2008**, *38*, 211–242.
- (188) Leman, L.; Orgel, L.; Ghadiri, M. R. Carbonyl Sulfide-Mediated Prebiotic Formation of Peptides. *Science* **2004**, *306*, 283–286.
- (189) E., O. L. Prebiotic Chemistry and the Origin of the RNA World. *Critical Reviews in Biochemistry and Molecular Biology* **2004**, *39*, 99–123.
- (190) Kanavarioti, A.; Monnard, P.-A.; Deamer, D. W. Eutectic Phases in Ice Facilitate Nonenzymatic Nucleic Acid Synthesis. *Astrobiology* **2001**, *1*, PMID: 12448990, 271–281.

- (191) Szymaszek, P.; Tomal, W.; Świergosz, T.; Kamińska-Borek, I.; Popielarz, R.; Ortyl, J. Review of quantitative and qualitative methods for monitoring photopolymerization reactions. *Polym. Chem.* **2023**, *14*, 1690–1717.
- (192) Aguirre-Soto, A.; Hwang, A. T.; Glugla, D.; Wydra, J. W.; McLeod, R. R.; Bowman, C. N.; Stansbury, J. W. Coupled UV–Vis/FT–NIR Spectroscopy for Kinetic Analysis of Multiple Reaction Steps in Polymerizations. *Macromolecules* **2015**, *48*, 6781–6790.
- (193) Quinebèche, S.; Navarro, C.; Gnanou, Y.; Fontanille, M. In situ mid-IR and UV–visible spectroscopies applied to the determination of kinetic parameters in the anionic copolymerization of styrene and isoprene. *Polymer* **2009**, *50*, 1351–1357.
- (194) Reis, M. H.; Leibfarth, F. A.; Pitet, L. M. Polymerizations in Continuous Flow: Recent Advances in the Synthesis of Diverse Polymeric Materials. *ACS Macro Letters* **2020**, *9*, PMID: 35638663, 123–133.
- (195) Kanao, M.; Matsuda, Y.; Sato, T. Characterization of Polymer Solutions Containing a Small Amount of Aggregates by Static and Dynamic Light Scattering. *Macromolecules* **2003**, *36*, 2093–2102.
- (196) Minton, A. P. Recent applications of light scattering measurement in the biological and biopharmaceutical sciences. *Analytical Biochemistry* **2016**, *501*, 4–22.
- (197) Bou, S.; Klymchenko, A. S.; Collot, M. Fluorescent labeling of biocompatible block copolymers: synthetic strategies and applications in bioimaging. *Mater. Adv.* **2021**, *2*, 3213–3233.
- (198) Morawetz, H. Studies of synthetic polymers carrying fluorescent labels. *Die Makromolekulare Chemie* **1985**, *13*, 67–74.
- (199) Uddin, M. A.; Yu, H.; Wang, L.; Naveed, K.-R.; Haq, F.; Amin, B. U.; Mehmood, S.; Nazir, A.; Xing, Y.; Shen, D. Recent progress in EPR study of spin labeled polymers and spin probed polymer systems. *Journal of Polymer Science* **2020**, *58*, 1924–1948.
- (200) Yu, L.; Lei, Y.; Ma, Y.; Liu, M.; Zheng, J.; Dan, D.; Gao, P. A Comprehensive Review of Fluorescence Correlation Spectroscopy. *Frontiers in Physics* **2021**, *9*, DOI: 10.3389/fphy.2021.644450.

- (201) Rajdev, P.; Ghosh, S. Fluorescence Resonance Energy Transfer (FRET): A Powerful Tool for Probing Amphiphilic Polymer Aggregates and Supramolecular Polymers. *The Journal of Physical Chemistry B* **2019**, *123*, 327–342.
- (202) Ye, R.; Sun, X.; Mao, X.; Alfonso, F. S.; Baral, S.; Liu, C.; Coates, G. W.; Chen, P. Optical sequencing of single synthetic polymers. *Nature Chemistry* **2024**, *16*, 210–217.
- (203) Domínguez, S. E.; Cangiotti, M.; Fattori, A.; Ääritalo, T.; Damlin, P.; Ottaviani, M. F.; Kvarnström, C. Effect of Spacer Length and Solvent on the Concentration-Driven Aggregation of Cationic Hydrogen-Bonding Donor Polythiophenes. *Langmuir* **2018**, *34*, PMID: 29783844, 7364–7378.
- (204) Barz, M.; Tarantola, M.; Fischer, K.; Schmidt, M.; Luxenhofer, R.; Janshoff, A.; Theato, P.; Zentel, R. From Defined Reactive Diblock Copolymers to Functional HPMA-Based Self-Assembled Nanoaggregates. *Biomacromolecules* **2008**, *9*, PMID: 18855476, 3114–3118.
- (205) Vekšli, Z.; Andreis, M.; Rakvin, B. ESR spectroscopy for the study of polymer heterogeneity. *Progress in Polymer Science* **2000**, *25*, 949–986.
- (206) Čulin, J.; Andreis, M.; Vekšli, Z.; Anžlovar, A.; Žigon, M. ESR-spin labelling study of semi-interpenetrating networks and polymer mixtures based on functionalized polyurethanes and polymethacrylates. *European Polymer Journal* **2005**, *41*, 1874–1882.
- (207) Ariciu, A. M.; Staicu, T.; Micutz, M.; Neacsu, M. V.; Ionita, P.; Tecuceanu, V.; Munteanu, C.; Ionita, G. Investigations on Carboxy Dibenzylidene Sorbitol Hydrogels Using EPR Spectroscopy. *Applied Magnetic Resonance* **2015**, *46*, 1395–1407.
- (208) Reichenwallner, J.; Thomas, A.; Nuhn, L.; Johann, T.; Meister, A.; Frey, H.; Hinderberger, D. Tunable dynamic hydrophobic attachment of guest molecules in amphiphilic core–shell polymers. *Polym. Chem.* **2016**, *7*, 5783–5798.
- (209) Xi, W.; Pattanayak, S.; Wang, C.; Fairbanks, B.; Gong, T.; Wagner, J.; Kloxin, C. J.; Bowman, C. N. Clickable Nucleic Acids: Sequence-Controlled Periodic Copolymer/Oligomer Synthesis by Orthogonal Thiol-X Reactions. *Angewandte Chemie International Edition* **2015**, *54*, 14462–14467.

- (210) Jeliński, T.; Bugalska, N.; Koszucka, K.; Przybyłek, M.; Cysewski, P. Solubility of sulfanilamide in binary solvents containing water: Measurements and prediction using Buchowski-Ksiazczak solubility model. *Journal of Molecular Liquids* **2020**, *319*, 114342.
- (211) Parkatzidis, K.; Wang, H. S.; Truong, N. P.; Anastasaki, A. Recent Developments and Future Challenges in Controlled Radical Polymerization: A 2020 Update. *Chem* **2020**, *6*, 1575–1588.
- (212) Matyjaszewski, K.; Spanswick, J. Controlled/living radical polymerization. *Materials Today* **2005**, *8*, 26–33.
- (213) Matyjaszewski, K.; Xia, J. Atom Transfer Radical Polymerization. *Chemical Reviews* **2001**, *101*, PMID: 11749397, 2921–2990.
- (214) Perrier, S. 50th Anniversary Perspective: RAFT Polymerization—A User Guide. *Macromolecules* **2017**, *50*, 7433–7447.
- (215) Nicolas, J.; Guillaneuf, Y.; Lefay, C.; Bertin, D.; Gimes, D.; Charleux, B. Nitroxide-mediated polymerization. *Progress in Polymer Science* **2013**, *38*, Topical Issue on Polymer Chemistry, 63–235.
- (216) Mocny, P.; Klok, H.-A. Complex polymer topologies and polymer—nanoparticle hybrid films prepared via surface-initiated controlled radical polymerization. *Progress in Polymer Science* **2020**, *100*, 101185.
- (217) Zoppe, J. O.; Ataman, N. C.; Mocny, P.; Wang, J.; Moraes, J.; Klok, H. A. Surface-Initiated Controlled Radical Polymerization: State-of-the-Art, Opportunities, and Challenges in Surface and Interface Engineering with Polymer Brushes. *Chemical Reviews* **2017**, *117*, 1105–1318.
- (218) Olivier, A.; Meyer, F.; Raquez, J.-M.; Damman, P.; Dubois, P. Surface-initiated controlled polymerization as a convenient method for designing functional polymer brushes: From self-assembled monolayers to patterned surfaces. *Progress in Polymer Science* **2012**, *37*, 157–181.
- (219) Hore, M. J. Polymers on nanoparticles: Structure & dynamics. *Soft Matter* **2019**, *15*, 1120–1134.
- (220) Kumar, S. K.; Ganesan, V.; Riggleman, R. A. Perspective: Outstanding theoretical questions in polymer-nanoparticle hybrids. *Journal of Chemical Physics* **2017**, *147*, DOI: 10.1063/1.4990501.

- (221) Gorman, C. B.; Petrie, R. J.; Genzer, J. Effect of substrate geometry on polymer molecular weight and polydispersity during surface-initiated polymerization. *Macromolecules* **2008**, *41*, 4856–4865.
- (222) Li, D.; Sheng, X.; Zhao, B. Environmentally responsive "hairy" nanoparticles: Mixed homopolymer brushes on silica nanoparticles synthesized by living radical polymerization techniques. *Journal of the American Chemical Society* **2005**, *127*, 6248–6256.
- (223) Gao, H.; Waechter, A.; Konstantinov, I. A.; Arturo, S. G.; Broadbelt, L. J. Application and comparison of derivative-free optimization algorithms to control and optimize free radical polymerization simulated using the kinetic Monte Carlo method. *Computers and Chemical Engineering* **2018**, *108*, 268–275.
- (224) Gao, H.; Oakley, L. H.; Konstantinov, I. A.; Arturo, S. G.; Broadbelt, L. J. Acceleration of Kinetic Monte Carlo Method for the Simulation of Free Radical Copolymerization through Scaling. *Industrial and Engineering Chemistry Research* **2015**, *54*, 11975–11985.
- (225) Genzer, J. In silico polymerization: Computer simulation of controlled radical polymerization in bulk and on flat surfaces. *Macromolecules* **2006**, *39*, 7157–7169.
- (226) Zapata-González, I.; Hutchinson, R. A.; Matyjaszewski, K.; Saldívar-Guerra, E.; Ortiz-Cisneros, J. Copolymer Composition Deviations from Mayo–Lewis Conventional Free Radical Behavior in Nitroxide Mediated Copolymerization. *Macromolecular Theory and Simulations* **2014**, *23*, 245–265.
- (227) Coleman, B. D.; Fox, T. G. A Multistate Mechanism for Homogeneous Ionic Polymerization. II. The Molecular Weight Distribution. *Journal of the American Chemical Society* **1963**, *85*, 1241–1244.
- (228) Thompson, A. P.; Aktulga, H. M.; Berger, R.; Bolintineanu, D. S.; Brown, W. M.; Crozier, P. S.; in 't Veld, P. J.; Kohlmeyer, A.; Moore, S. G.; Nguyen, T. D.; Shan, R.; Stevens, M. J.; Tranchida, J.; Trott, C.; Plimpton, S. J. LAMMPS - A Flexible Simulation Tool for Particle-Based Materials Modeling at the Atomic, Meso, and Continuum Scales. *Comp. Phys. Comm.* **2022**, *271*, 108171.
- (229) Campello, R. J. G. B.; Moulavi, D.; Sander, J. In *Advances in Knowledge Discovery and Data Mining*, ed. by Pei, J.; Tseng, V. S.; Cao, L.; Motoda, H.; Xu, G., Springer Berlin Heidelberg: Berlin, Heidelberg, 2013, pp 160–172.