



Soft Matter

**Selective solvent conditions influence sequence development and supramolecular assembly in step-growth copolymerization**

Journal:	<i>Soft Matter</i>
Manuscript ID	SM-ART-11-2021-001571.R1
Article Type:	Paper
Date Submitted by the Author:	24-Nov-2021
Complete List of Authors:	Hamblin, Ryan; University of Virginia, Chemistry Nguyen, Nhu; University of Virginia, Chemistry DuBay, Kateri; University of Virginia, Chemistry

SCHOLARONE™  
Manuscripts

Cite this: DOI: 00.0000/xxxxxxxxxx

# Selective solvent conditions influence sequence development and supramolecular assembly in step-growth copolymerization<sup>†</sup>

Ryan L. Hamblin<sup>a</sup>, Nhu Q. Nguyen<sup>a</sup>, and Kateri H. DuBay<sup>\*a</sup>

Received Date

Accepted Date

DOI: 00.0000/xxxxxxxxxx

Sequence control in synthetic copolymers remains a tantalizing objective in polymer science due to the influence of sequence on material properties and self-organization. A greater understanding of sequence development throughout the polymerization process will aid the design of simple, generalizable methods to control sequence and tune supramolecular assembly. In previous simulations of solution-based step-growth copolymerizations, we have shown that weak, non-bonding attractions between monomers of the same type can produce a microphase separation among the lengthening nascent oligomers and thereby alter sequence. This work explores the 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 find 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. Our results demonstrate the impact of differing A-B monomer-solvent affinities on sequence development in solution-based copolymerizations and are of consequence to the informed design of synthetic methods for sequence controlled amphiphilic copolymers and their aggregates.

## 1 Introduction

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>1–3</sup> Both the material properties and macromolecular morphologies of synthetic copolymers are intrinsically dependent upon their sequences and assembly behavior.<sup>4–8</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.<sup>9–11</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 ap-

proaches.<sup>12,13</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>1,2,14</sup> Our previous work explored such sequence influencing effects in a simplified model of a solvated step-growth polymerization among two monomer types. We demonstrated that the presence of even relatively weak non-bonding attractions between monomers can induce an emergent microphase separation as the reaction proceeds.<sup>14</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>15</sup>

This type of local concentration enrichment has previously been reported in both experimental and computational observations of the "bootstrap" effect,<sup>13,16–21</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 stud-

<sup>a</sup> University of Virginia, Department of Chemistry, McCormick Road, PO Box 400319, Charlottesville, VA 22903-4319, USA.

E-mail: [dubay@virginia.edu](mailto:dubay@virginia.edu)

<sup>†</sup> Electronic Supplementary Information (ESI) available. See DOI: 10.1039/cXsm00000x/

ied 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>22–24</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>22,25–28</sup>

In our past studies, 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.<sup>14,15</sup> 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>29–31</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>22,23</sup>

In this work we expand upon our prior studies to consider the case of a selective solvent, utilizing an extension of our coarse-grained, reactive model for step-growth polymerization employed in previous work. 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 Section 2. In Section 3.1 we examine the kinetic behaviors of copolymerization under such asymmetric attractions. In Section 3.2 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 3.3 and discuss the aggregation behavior of the resultant oligomers in Section 3.4. 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.

## 2 Methods

**Monomer structure and intermolecular interactions.** Our model is a coarse-grained, reactive representation of a step-growth copolymerization, first developed by Zhang and DuBay<sup>14</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>32,33</sup> A single bifunctional monomer is represented as a simplified bead, which represents the repeat unit and determines the identity of the monomer, denoted **A** or **B** to represent the two distinct monomer types taking part in the reaction. Each bead is composed of three particles, a central particle of type **1** and two additional interacting particles of type **2**, as shown schematically in Fig. 1a. Two harmonic bonds control the inter-particle spatial

separation and an angular harmonic potential controls the internal **2-1-2** angle, angle  $\alpha$  in Fig. 1b, 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.

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>14</sup>:

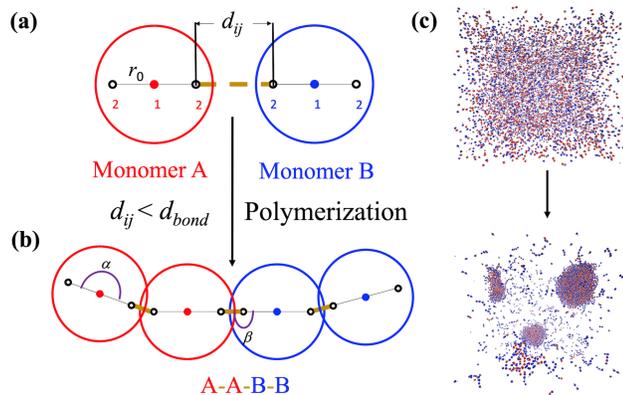
$$E_{\text{LJ}}(\mathbf{1}, \mathbf{1}') = \begin{cases} 4\epsilon_{\text{att}}(\mathbf{1}, \mathbf{1}') \left[ \left( \frac{\sigma}{r(\mathbf{1}, \mathbf{1}')} \right)^{12} - \left( \frac{\sigma}{r(\mathbf{1}, \mathbf{1}')} \right)^6 \right] & r_0 \leq r(\mathbf{1}, \mathbf{1}') < 2.5\sigma \\ 4\epsilon_{\text{rep}}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

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  $\epsilon_{\text{rep}}$ .

The type **2** interaction particles of each monomer bead represent the reactive functional groups of the monomer, and their non-bonded interactions take the form of a purely repulsive soft potential. This 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, as shown in the transition between Fig. 1a and Fig. 1b. Upon bond formation, all non-bonded interactions between directly bound monomers are switched off.

**Simulation progression and bond formation.** The simulation begins with equal numbers of **A** and **B** monomers distributed randomly throughout the simulation volume and is first allowed to equilibrate with all attractive interactions and bond formation turned off. Time evolution occurs according to the Langevin equation, which simultaneously controls the temperature of the system and provides the characteristics of viscous interactions between monomers and an implicit solvent. 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. 1c shows the initial and final system state for a representative simulation under these conditions.

**Asymmetric effective inter-monomer interactions.** This work seeks to explore copolymerizations in which monomers have asymmetric non-bonded interaction strengths, by impos-



**Fig. 1 Polymerization model schematic.** (a) Each monomer is composed of a central particle (type 1) and two binding particles (type 2) with their internal structure governed by harmonic bonds. Monomers also interact via nonbonded interactions which depend on the monomer identity, either A or B. When a type 2 particle on one monomer comes within a distance  $d_{bond}$  of a type 2 particle on another monomer, an irreversible bond is formed between them. (b) Through sequential bond formation events, oligomers form. The intra-monomer angle,  $\alpha$  (1–2–1), and the inter-monomer angle,  $\beta$  (2–1–1'), are also shown here. Polymer chain stiffness, and thus persistence length,  $l_p$ , are controlled by tuning the angular harmonic potential associated with the intra-monomer angle,  $\alpha$ . (c) Polymerization proceeds from randomly distributed, well-dissolved, monomers and forms more complex structures as they polymerize.

ing 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>34,35</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>14</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 S1 in the SI. 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>15</sup> Three separate trials were performed

for all sets of simulation parameters. A more detailed description of the model is available in the SI.

## 3 Results

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

Our first aim was 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 occurs, which alters the solution structure and produces unconventional kinetic features. Fig. 2a 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>36</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>36</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  $X_n$  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 the solution environment during the transition, as quantified by the radial distribution function between different monomer species,  $g_{ij}(r)$ , shown in Fig. 2b&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. 2b. 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. 2c. 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 Section S4 in SI 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,

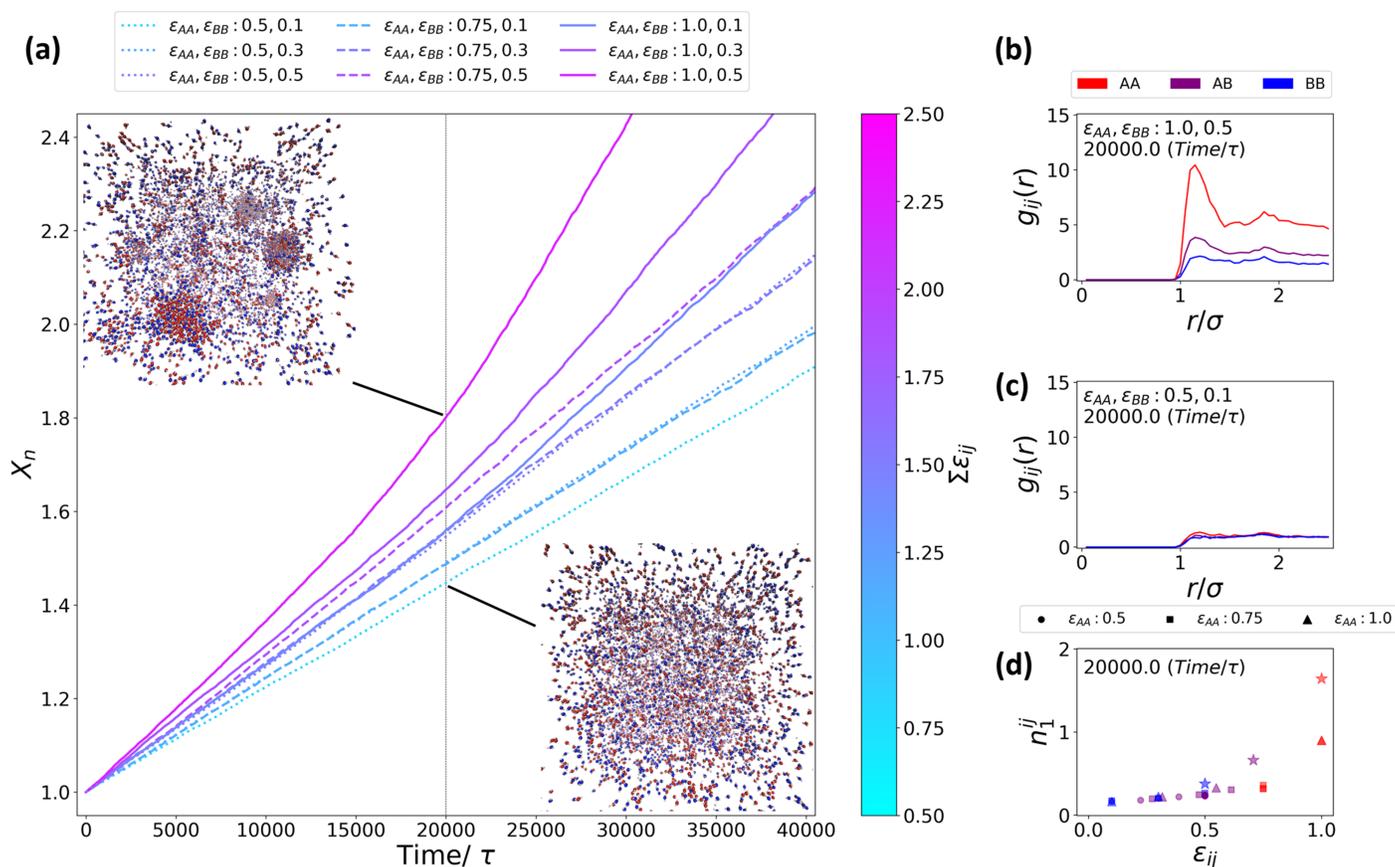
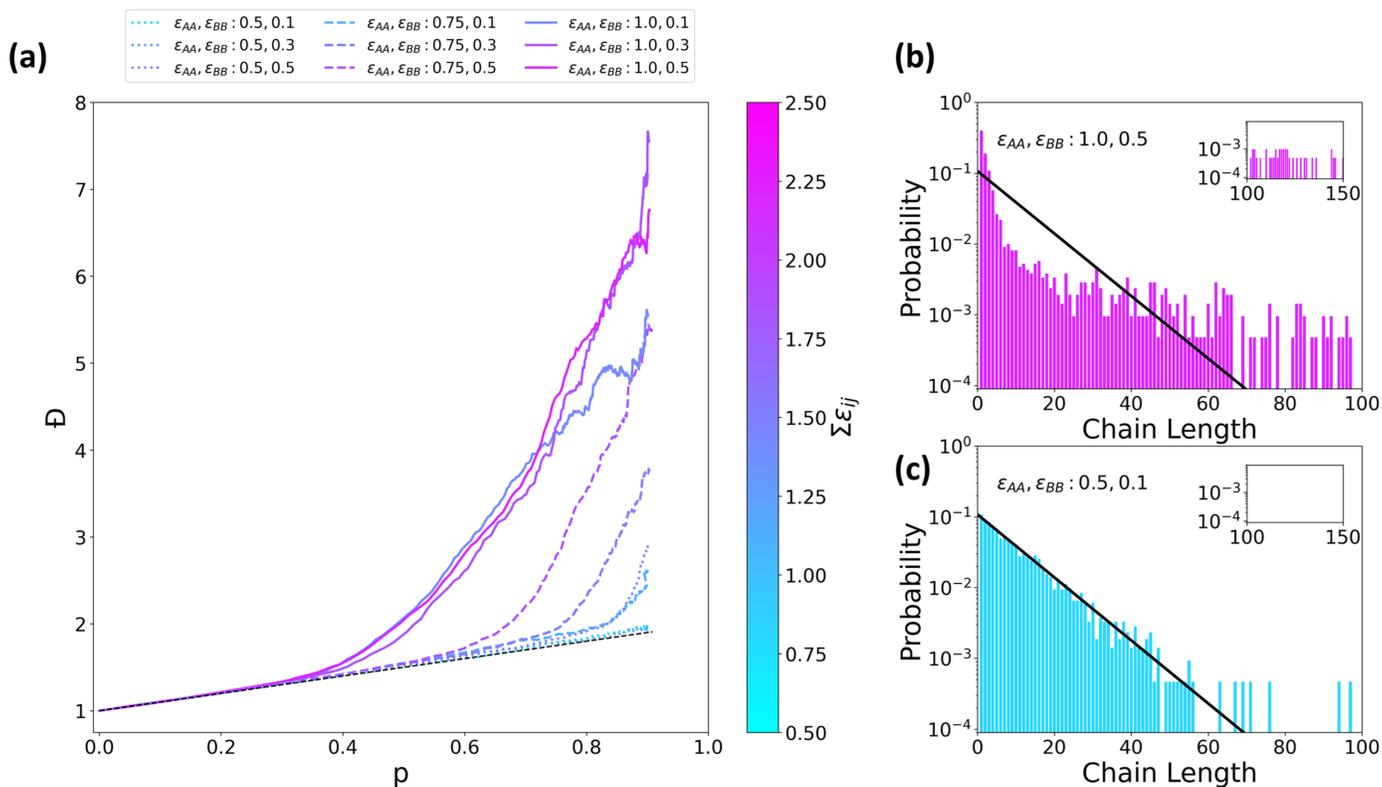


Fig. 2 **Local monomer density and polymerization kinetics.** (a) Number averaged degree of polymerization  $X_n$  vs. simulation time for each combination of attractive interaction strengths explored. Dotted lines indicate simulations with  $\epsilon_{AA} = 0.5k_B T$ , dashed lines indicate  $\epsilon_{AA} = 0.75k_B T$ , 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  $ij = AA$ , blue for  $ij = BB$ , and purple for  $ij = AB$  or  $ij = BA$ . (d) First coordination number,  $n_1^{ij}$ , vs.  $\epsilon_{ij}$  is shown by monomer type pairs for each attraction strength explored.

are presented as a function of the governing attraction strength,  $\epsilon_{ij}$ , in Fig. 2d. 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>36</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 re-

acting 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>36</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>36</sup>. Figure 3a 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. 2a, 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. 2a, 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



**Fig. 3 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  $\varepsilon_{AA} = 0.5k_B T$ , dashed lines indicate  $\varepsilon_{AA} = 0.75k_B T$ , and solid lines indicate simulations with  $\varepsilon_{AA} = k_B T$ . The coloration indicates the total attraction strength  $\Sigma \varepsilon_{ij}$  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)  $\varepsilon_{AA} = 0.5k_B T$ ,  $\varepsilon_{BB} = 0.1k_B T$  and (c)  $\varepsilon_{AA} = k_B T$ ,  $\varepsilon_{BB} = 0.5k_B T$ . 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.

over the full polymerization are available in Figures S5 and S6 in the SI.

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. 3b&c. For the lowest attraction strength in Fig. 3c, the expected Flory-Schulz distribution (solid black line) matches the observed chain length distribution. However, for the higher attraction strength shown in Fig. 3b, the observed chain length distribution is noticeably shifted towards longer chain lengths. Overall, Fig. 3 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  $\mathcal{D}$ .

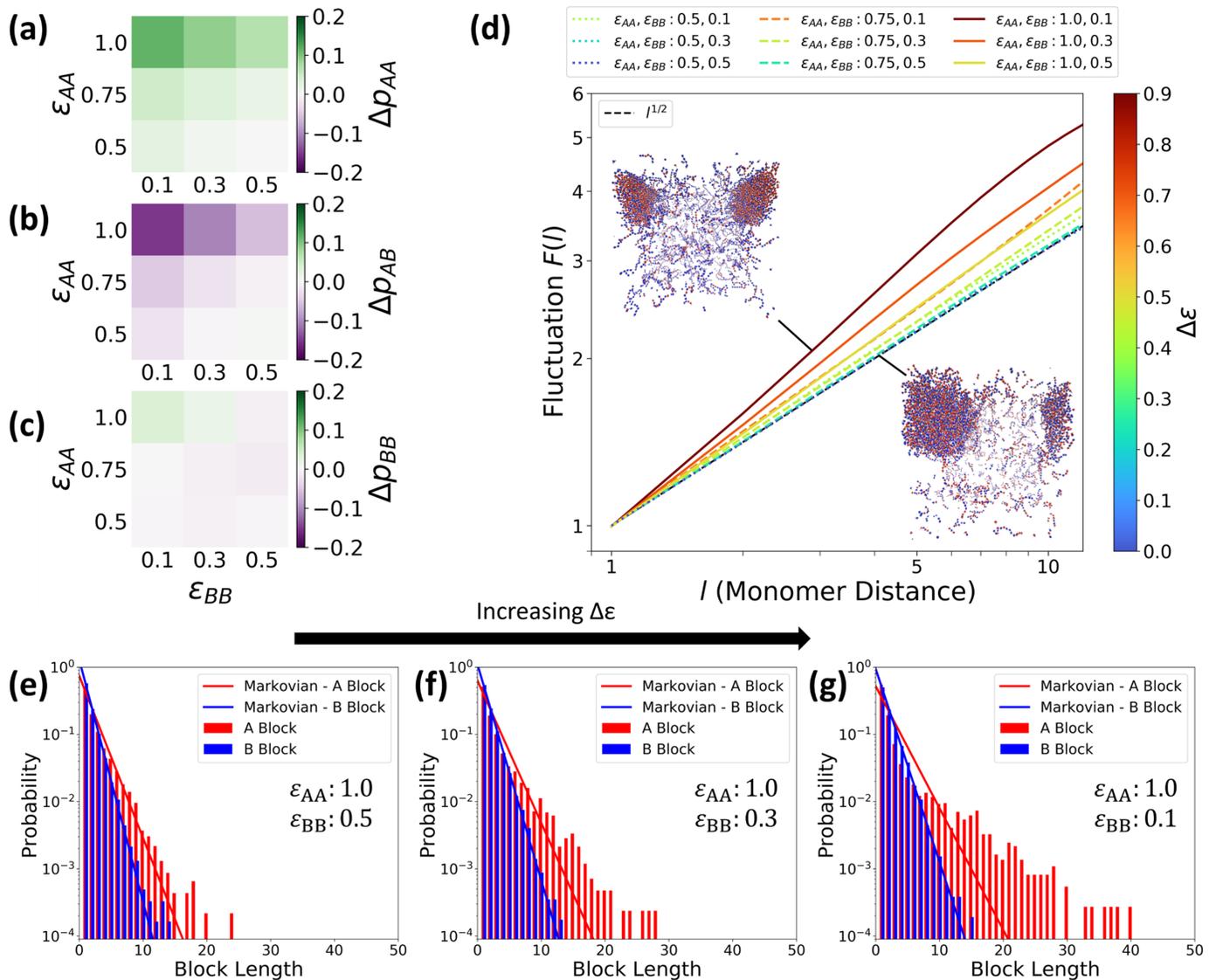
### 3.2 Oligomer sequences depend on solvent selectivity.

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. 4a-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 for  $p_{ij}$  change throughout the reaction as the sequences develop, as can be seen in the SI Figures S5 and S6.

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,



**Fig. 4 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 (see SI for details) 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  $p = 0.9$ .

$\Delta\epsilon \equiv \epsilon_{AA} - \epsilon_{BB}$ , 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>14,15,37</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. 4d 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.

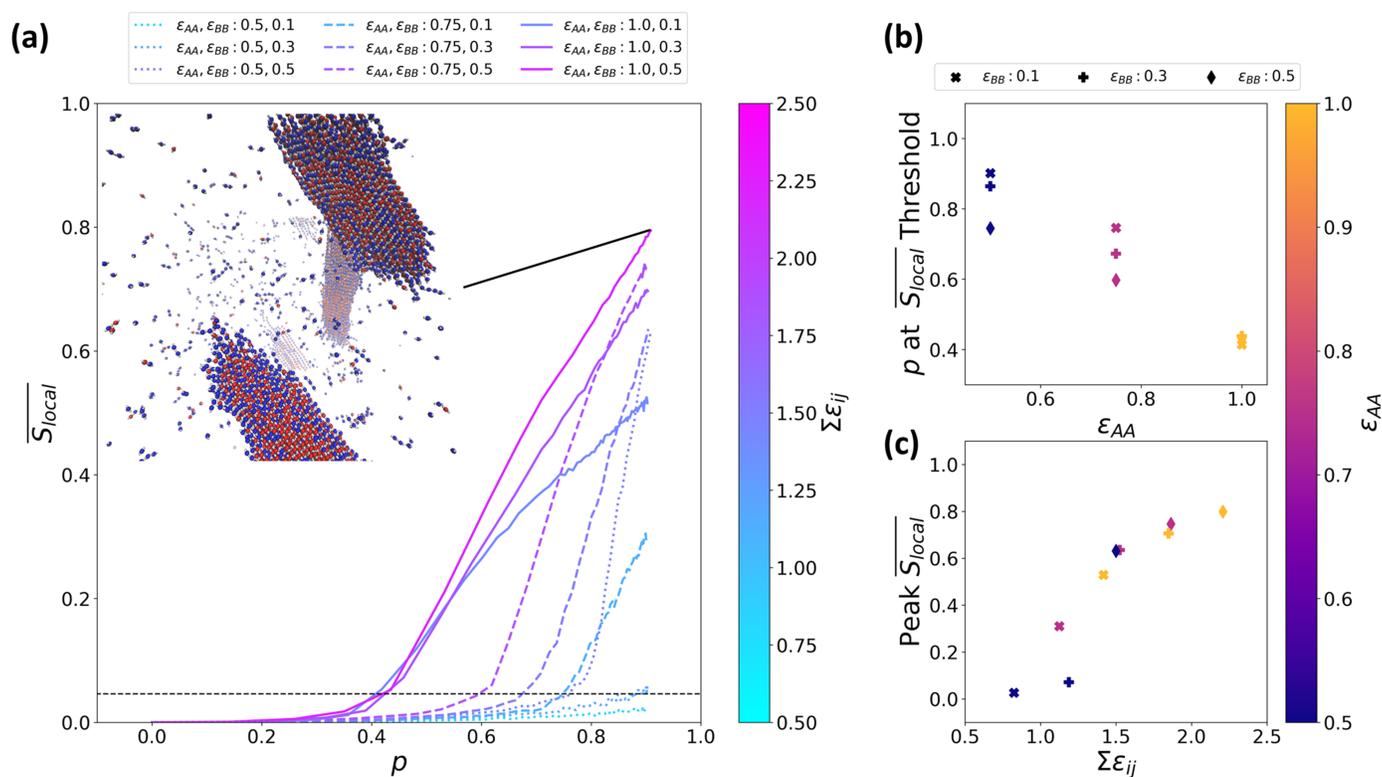
The result of these long-range correlations in sequence may be seen in the length distributions of contiguous blocks of either **A** or **B** monomers within the oligomers. Fig. 4e-g shows the block length distributions for a fixed  $\epsilon_{AA} = k_B T$  and an increasing  $\Delta\epsilon$ . Similar 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 Section S5 in the SI for details). 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. 3b). 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. 4d 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.

### 3.3 Oligomer persistence length impacts both sequence and aggregate structure.

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>15</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 Section S6 in the SI 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. 5a 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  $\Sigma\epsilon_{ij} \gtrsim 1.2k_B T$ , where  $\overline{S}_{\text{local}}$  reaches values  $\geq 0.5$ , indicating significant alignment amongst neighboring chains. For the highest value of  $\Sigma\epsilon_{ij}$  at  $p = 0.9$ , shown in the inset of Fig. 5a, 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. S3 in SI), 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. 5a. 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. 5b&c show the trend in these properties as a function of  $\epsilon_{AA}$  and  $\Sigma\epsilon_{ij}$  respectively. Fig. 5b shows that the point at which  $\overline{S}_{\text{local}}$  exceeds the threshold established by the flexible chain case is predominantly controlled by  $\epsilon_{AA}$ . For  $\epsilon_{AA} = k_B T$  in particular, this onset of ordering is largely independent of the value of  $\epsilon_{BB}$  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**



**Fig. 5 Localized chain alignments in stiff-chain oligomers.** (a) The local nematic ordering parameter,  $\overline{S}_{local}$ , 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,  $\Sigma \epsilon_{ij}$ , for the simulation. The dashed horizontal black line indicates the highest value for  $\overline{S}_{local}$  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  $\overline{S}_{local}$  exceeds the threshold value is shown for each parameter set as a function of  $\epsilon_{AA}$ . (c) The peak value of  $\overline{S}_{local}$  reached for each simulation parameter set as a function of  $\Sigma \epsilon_{ij}$ . 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.

monomers into the polymer phase is significantly less than that of A (see Fig. S3). The  $\overline{S}_{local}$  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  $\overline{S}_{local}$  values at  $p = 0.9$  are thus predominantly controlled by  $\Sigma \epsilon_{ij}$  as seen in Fig. 5c. 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.

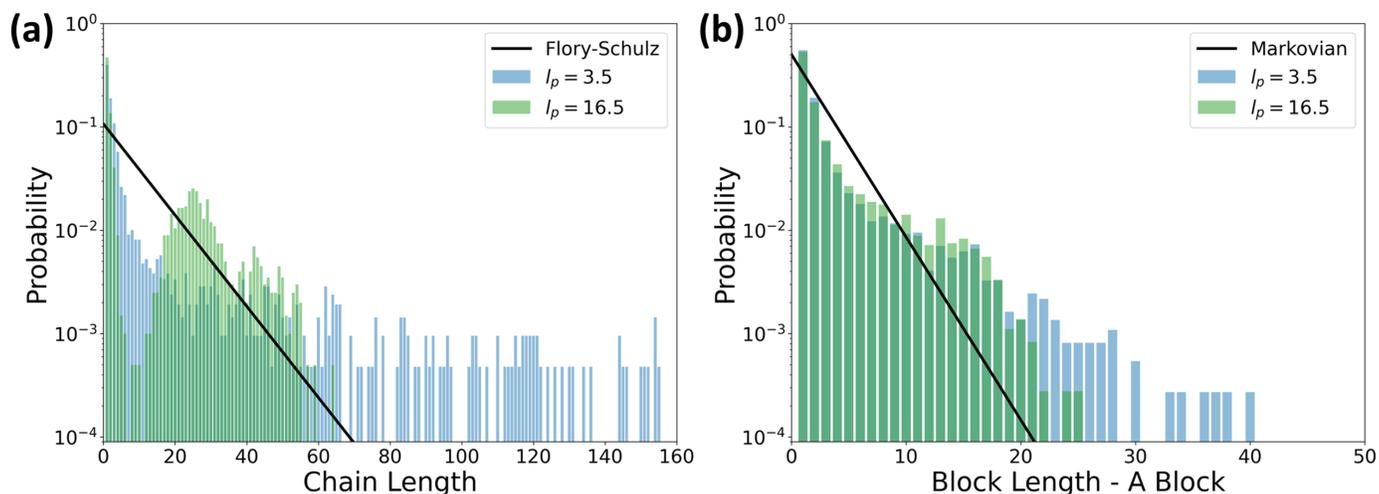
The impact of chain stiffness and nematic alignment on chain and block lengths is significant. Fig. 6 shows the chain and block length distributions at  $p = 0.9$  for systems with two different persistence lengths but the same set of attractive interactions. In both cases, attractive interactions strong 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 persis-

tence length and associated orientational ordering. These results are similar to what was previously seen in the case of symmetric attractions.<sup>15</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). 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.

### 3.4 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. 2 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>36</sup> Af-



**Fig. 6 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  $\varepsilon_{AA} = k_B T$ ,  $\varepsilon_{BB} = 0.5k_B T$ . 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  $\varepsilon_{AA} = k_B T$ ,  $\varepsilon_{BB} = 0.1k_B T$  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.

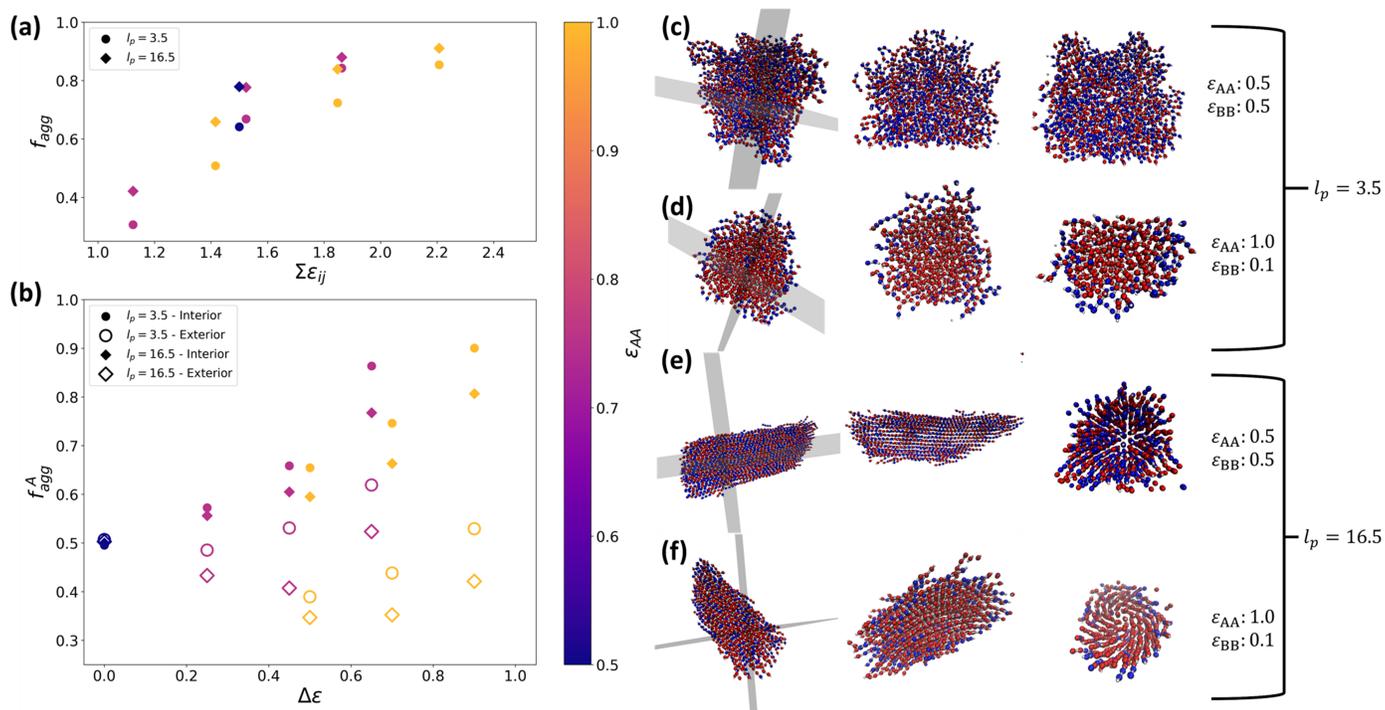
ter 90% of possible bonds in the simulation have formed, this aggregated phase has developed into clearly distinct structures of clustered oligomers. The nematic alignment observed at high persistence lengths is fundamentally driven by the formation of dense oligomer aggregates. To identify properties 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.

Fig. 7 shows the results of the application of these local neighbor criteria to our system at  $p = 0.9$ . Notably, for systems with  $\varepsilon_{AA} = 0.5k_B T$  and  $\varepsilon_{BB} < 0.5k_B T$ , we do not find regions for which  $n^{\text{local}} \geq 12$ . At these attractions, the reaction kinetics (Fig. 2a) and dispersities (Fig. 3) are as expected for an ideal step-growth process, the  $g(r)$  function shows no concentration enrichment early in the reaction (see Fig. 2c), 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. 7a, the total fraction of all monomers in

the aggregated phase,  $f_{\text{agg}}$ , is shown as a function of the total attraction strength,  $\Sigma\varepsilon_{ij}$ . As with the aggregation and the associated rate change that occurs earlier in the polymerization (see Fig. 2), 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,  $f_{\text{agg}}^A$ , within each region. The results for  $f_{\text{agg}}^A$  are shown in Fig. 7b for all simulation parameters where aggregates form. For the symmetric attraction case,  $\Delta\varepsilon = 0$ , both monomer species are equally incorporated into the interior and exterior regions of the aggregates. Representative aggregate structures under symmetric attractions (Fig. 7c&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. 7d&f.

The impact of increased chain stiffness on the formation and structure of aggregates is also apparent from the results in Fig. 7. Increased chain stiffness has been shown to influence the phase behavior of polymers, and greater stiffness further promotes the



**Fig. 7 Aggregate compositions and representative structures.** (a)  $f_{agg}$ , the fraction of all monomers in the system that are located within an aggregate (as defined by  $n^{local} \geq 12$ ), is shown as a function of the total attraction strength,  $\Sigma\epsilon_{ij}$ . (b)  $f_{agg}^A$ , the fraction of A monomers within an aggregate, is plotted as a function of  $\Delta\epsilon$ . Monomers with  $n^{local} \geq 36$  were identified as “interior” (filled markers), and monomers with  $12 \leq n^{local} < 36$  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  $p = 0.9$ .

formation of an aggregated phase.<sup>38–42</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. 7a). 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. 7b. This reduction of interior  $f_{agg}^A$  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. 7a) 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. 7c&d), and stiff chains forming elongated, rod-like aggregates (Fig. 7e&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 Discussion

These results extend our previous observations of kinetic alteration and sequence biasing effects in step-growth polymerizations<sup>14,15</sup> to cases where effective intermonomer attractions are asymmetric, such as would arise when solvent affinities differ between reactants. These conditions relate to a broad class of real-world polymerizations where such a difference in reactant affinities exist.<sup>29–31,43</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>13,16,17</sup> and for the reaction rate increases in PISA formulations.<sup>22,25,26</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>22,44,45</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>46–48</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>42,49–51</sup>

More generally, the observed assembly process belongs to the broad class of non-classical self-assembly pathways.<sup>52–56</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>57–60</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>36</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>61,62</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>63,64</sup> Other polymerization induced microphase separations have been shown to promote desirable material properties as well,<sup>65,66</sup> and there is continued interest in the use of solvent interactions to tune supramolecular polymer structures.<sup>43</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,  $\epsilon_{AA} = k_B T$  (which is  $\sim 2.5$  kJ/mol at 300 K), are readily accessible interaction energies for a variety of supramolecular interactions,<sup>67–69</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>70–74</sup> Interestingly, according to Flory-Huggins theory,<sup>36</sup> even the milder attractions of  $\sim 0.3k_B T$ , 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.

## 5 Conclusion

We have utilized a reactive, coarse-grained model to explore 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 syn-

thetic 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.

### Author Contributions

**Ryan L. Hamblin:** Investigation, Software, Formal analysis, Visualization, Writing - original draft. **Nhu Q. Nguyen:** Data curation, Visualization, Writing - review & editing. **Kateri H. DuBay:** Conceptualization, Methodology, Supervision, Formal analysis, Writing - original draft.

### Conflicts of interest

There are no conflicts to declare.

### Acknowledgements

This material is based upon work supported by the National Science Foundation under grant no. 1848009. The authors would like to thank Zhongmin Zhang and Lutz Maibaum for helpful discussions. The authors also acknowledge Research Computing at The University of Virginia (<https://rc.virginia.edu>) for providing computational resources and technical support that have contributed to the results reported within this publication.

### References

- J.-F. Lutz, M. Ouchi, D. R. Liu and M. Sawamoto, *Science*, 2013, **341**, 1238149.
- J.-F. Lutz, *Macromolecular Rapid Communications*, 2017, **38**, 1700582.
- J. De Neve, J. J. Haven, L. Maes and T. Junkers, *Polym. Chem.*, 2018, **9**, 4692–4705.
- K. Matyjaszewski, M. J. Ziegler, S. V. Arehart, D. Greszta and T. Pakula, *Journal of Physical Organic Chemistry*, 2000, **13**, 775–786.
- S. Zhang, N. E. Bauer, I. Y. Kanal, W. You, G. R. Hutchison and T. Y. Meyer, *Macromolecules*, 2017, **50**, 151–161.
- L. A. Strickland, C. K. Hall and J. Genzer, *Macromolecules*, 2009, **42**, 9063–9071.
- R. Malik, C. K. Hall and J. Genzer, *Soft Matter*, 2011, **7**, 10620–10630.
- A. M. Rumyantsev, A. Johnner and J. J. de Pablo, *ACS Macro Letters*, 2021, **10**, 1048–1054.
- M. M. Mok, J. Kim and J. M. Torkelson, *Journal of Polymer Science Part B: Polymer Physics*, 2008, **46**, 48–58.
- X.-F. Li, F. P. V. Paoloni, E. A. Weiber, Z.-H. Jiang and P. Janasch, *Macromolecules*, 2012, **45**, 1447–1459.
- T. K. Lytle, L.-W. Chang, N. Markiewicz, S. L. Perry and C. E. Sing, *ACS Central Science*, 2019, **5**, 709–718.
- S. I. Kuchanov, M. L. Keshtov, P. G. Halatur, V. A. Vasnev, S. V. Vinogradova and V. V. Korshak, *Die Makromolekulare Chemie*, 1983, **184**, 105–111.
- O. V. Borisova, M. Y. Zaremski, O. V. Borisov and L. Billon, *Polymer Science - Series B*, 2013, **55**, 573–576.
- Z. Zhang and K. H. DuBay, *Macromolecules*, 2019, **52**, 5480–5490.
- Z. Zhang and K. H. DuBay, *The Journal of Physical Chemistry B*, 2021, **125**, 3426–3437.
- H. J. Harwood, *Makromolekulare Chemie. Macromolecular Symposia*, 1987, pp. 331–354.
- M. L. Coote, T. P. Davis, B. Klumperman and M. J. Monteiro, *Journal of Macromolecular Science, Part C*, 1998, **38**, 567–593.
- O. A. Kazantsev, D. M. Kamorin, D. V. Orekhov and A. P. Sivokhin, *Designed Monomers and Polymers*, 2015, **18**, 378–384.
- A. A. Gavrilov and A. V. Chertovich, *Macromolecules*, 2017, **50**, 4677–4685.
- S.-M. Kong, H. Liu, Y.-H. Xue, X.-L. Liu, X.-X. Jia and F.-C. Cui, *Phys. Chem. Chem. Phys.*, 2018, **20**, 24379–24388.
- J. Hao, F. An, C. Lu and Y. Liu, *Journal of Macromolecular Science, Part A*, 2019, **56**, 1012–1021.
- N. J. Penfold, J. Yeow, C. Boyer and S. P. Armes, *ACS Macro Letters*, 2019, **8**, 1029–1054.
- S. L. Canning, G. N. Smith and S. P. Armes, *Macromolecules*, 2016, **49**, 1985–2001.
- B. Charleux, G. Delaittre, J. Rieger and F. D'Agosto, *Macromolecules*, 2012, **45**, 6753–6765.
- I. Chaduc, M. Girod, R. Antoine, B. Charleux, F. D'Agosto and M. Lansalot, *Macromolecules*, 2012, **45**, 5881–5893.
- A. Blanazs, J. Madsen, G. Battaglia, A. J. Ryan and S. P. Armes, *Journal of the American Chemical Society*, 2011, **133**, 16581–16587.
- W. J. Zhang, C. Y. Hong and C. Y. Pan, *Macromolecular Rapid Communications*, 2019, **40**, 1–10.
- X. Wang and Z. An, *Macromolecular Rapid Communications*, 2019, **40**, 1–14.
- D. Heinz, E. Amado and J. Kressler, *Polymers*, 2018, **10**.
- R. Matsidik, H. Komber and M. Sommer, *ACS Macro Letters*, 2015, **4**, 1346–1350.
- K. Plochocka, *Journal of Macromolecular Science, Part C*, 1981, **20**, 67–148.
- P. E. Rouse, *The Journal of Chemical Physics*, 1953, **21**, 1272–1280.
- D. Fritz, K. Koschke, V. A. Harmandaris, N. F. A. van der Vegt and K. Kremer, *Phys. Chem. Chem. Phys.*, 2011, **13**, 10412–10420.
- W. B. Brown and G. Gee, *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 1957, **250**, 175–220.
- C. L. Kong, *The Journal of Chemical Physics*, 1973, **59**, 2464–2467.
- P. J. Flory, *Principles of Polymer Chemistry*, Cornell University Press, 1953.
- C.-K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, F. Sciortino, M. Simons and H. Stanley, *Nature*, 1992, **356**, 168–170.
- P. J. Flory, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 1956, **234**, 60–73.

- 39 E. DiMarzio and J. Gibbs, *The Journal of Chemical Physics*, 1958, **28**, 807–813.
- 40 G. H. Fredrickson, A. J. Liu and F. S. Bates, *Macromolecules*, 1994, **27**, 2503–2511.
- 41 J. Zierenberg and W. Janke, *EPL (Europhysics Letters)*, 2015, **109**, 28002.
- 42 J. Midya, S. A. Egorov, K. Binder and A. Nikoubashman, *The Journal of Chemical Physics*, 2019, **151**, 034902.
- 43 M. F. J. Mabesoone, A. R. A. Palmans and E. W. Meijer, *Journal of the American Chemical Society*, 2020, **142**, 19781–19798.
- 44 R. A. Ramli, W. A. Laftah and S. Hashim, *RSC Adv.*, 2013, **3**, 15543–15565.
- 45 P. A. Lovell and F. J. Schork, *Biomacromolecules*, 2020, **21**, 4396–4441.
- 46 V. S. Kravchenko and I. I. Potemkin, *The Journal of Physical Chemistry B*, 2016, **120**, 12211–12217.
- 47 V. S. Kravchenko, V. Abetz and I. I. Potemkin, *Polymer*, 2021, **235**, 124288.
- 48 G. Pandav, V. Pryamitsyn, K. C. Gallow, Y.-L. Loo, J. Genzer and V. Ganesan, *Soft Matter*, 2012, **8**, 6471–6482.
- 49 Y. M. Gross and S. Ludwigs, *Synthetic Metals*, 2019, **253**, 73–87.
- 50 P. I. Kos, V. A. Ivanov and A. V. Chertovich, *Soft Matter*, 2021, **17**, 2392–2403.
- 51 K. E. Polovnikov and I. I. Potemkin, *The Journal of Physical Chemistry B*, 2017, **121**, 10180–10189.
- 52 D. Kashchiev, P. G. Vekilov and A. B. Kolomeisky, *The Journal of Chemical Physics*, 2005, **122**, 244706.
- 53 D. Erdemir, A. Y. Lee and A. S. Myerson, *Accounts of Chemical Research*, 2009, **42**, 621–629.
- 54 J. J. D. Yoreo, P. U. P. A. Gilbert, N. A. J. M. Sommerdijk, R. L. Penn, S. Whitelam, D. Joester, H. Zhang, J. D. Rimer, A. Navrotsky, J. F. Banfield, A. F. Wallace, F. M. Michel, F. C. Meldrum, H. Cölfen and P. M. Dove, *Science*, 2015, **349**, aaa6760.
- 55 J. L. Zryd and W. R. Burghardt, *Journal of Applied Polymer Science*, 1995, **57**, 1525–1537.
- 56 O. Galkin and P. G. Vekilov, *Proceedings of the National Academy of Sciences*, 2000, **97**, 6277–6281.
- 57 W. Hu, D. Frenkel and V. B. F. Mathot, *Macromolecules*, 2002, **35**, 7172–7174.
- 58 W. Hu and D. Frenkel, in *Polymer Crystallization Driven by Anisotropic Interactions*, ed. G. Allegra, Springer Berlin Heidelberg, Berlin, Heidelberg, 2005, pp. 1–35.
- 59 T. K. Haxton and S. Whitelam, *Soft Matter*, 2012, **8**, 3558–3562.
- 60 T. K. Haxton, L. O. Hedges and S. Whitelam, *Soft Matter*, 2015, **11**, 9307–9320.
- 61 L. Wang, Y. Ding, Q. Liu, Q. Zhao, X. Dai, X. Lu and Y. Cai, *ACS Macro Letters*, 2019, **8**, 623–628.
- 62 S. Xu, T. Zhang, R. P. Kuchel, J. Yeow and C. Boyer, *Macromolecular Rapid Communications*, 2020, **41**, 1900493.
- 63 J. Lesage de la Haye, X. Zhang, I. Chaduc, F. Brunel, M. Lansalot and F. D'Agosto, *Angewandte Chemie International Edition*, 2016, **55**, 3739–3743.
- 64 F. Brunel, J. Lesage De La Haye, M. Lansalot and F. D'Agosto, *Journal of Physical Chemistry B*, 2019, **123**, 6609–6617.
- 65 K. A. Masser, E. D. Bain, F. L. Beyer, A. M. Savage, J. H. Yu and J. L. Lenhart, *Polymer*, 2016, **103**, 337–346.
- 66 M. J. Stevens, *Journal of Chemical Physics*, 2021, **155**, 1–9.
- 67 F. Biedermann and H.-J. Schneider, *Chemical Reviews*, 2016, **116**, 5216–5300.
- 68 M. D. Driver, M. J. Williamson, J. Cook and C. A. Hunter, *Chem. Sci.*, 2020, **11**, 4456–4466.
- 69 C. A. Hunter, *Angewandte Chemie International Edition*, 2004, **43**, 5310–5324.
- 70 R. Ramachandran, G. Beaucage, A. S. Kulkarni, D. McFaddin, J. Merrick-Mack and V. Galiatsatos, *Macromolecules*, 2008, **41**, 9802–9806.
- 71 C. L. Gettinger, A. J. Heeger, J. M. Drake and D. J. Pine, *The Journal of Chemical Physics*, 1994, **101**, 1673–1678.
- 72 S. Brinkers, H. R. Dietrich, F. H. de Groote, I. T. Young and B. Rieger, *The Journal of chemical physics*, 2009, **130**, 06B607.
- 73 R. H. Walters and R. M. Murphy, *Journal of Molecular Biology*, 2009, **393**, 978–992.
- 74 A. Chin, D. Toptygin, W. Elam, T. Schrank and V. Hilser, *Biophysical Journal*, 2016, **110**, 362–371.