A Patchy Particle Model of the Hierarchical Self-Assembly of π-Conjugated Optoelectronic Peptides

Rachael A. Mansbach[†,‡] and Andrew L. Ferguson[*,¶,§,†,‖]

†*Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, IL 61801, USA.*

‡*Present address: Los Alamos National Laboratory, Theoretical Biology and Biophysics Group, Los Alamos, NM 87545, USA.*

¶*Department of Materials Science and Engineering, University of Illinois at Urbana-Champaign, 1304 W Green Street, Urbana, IL 61801, USA.*

§*Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, IL 61801, USA.*

‖*Present address: Institute for Molecular Engineering, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA.*

E-mail: andrewferguson@uchicago.edu

Phone: (773) 702-3018

Abstract

Self-assembling peptides containing aromatic groups are an attractive target for bioelectronic materials design due to their ease of manufacture, biocompatibility, aqueous solubility, and chemical tunability. Microscopic understanding of the properties that control assembly is a prerequisite for rational design. In this work, we study the assembly of a family of DXXX-Π-XXXD oligopeptides possessing a π-conjugated core flanked by Asp-terminated tetrapeptide wings that display pH-triggered assembly into supramolecular aggregates. We develop a coarse-grained patchy particle model to conduct molecular dynamics simulations of the assembly of ten thousand oligopeptides over hundreds of nanometers and hundreds of microseconds. We study the effects of core and side chain interaction strength and side chain steric volume upon the morphology and kinetics of assembly. By characterizing the rate and fractal dimension of hierarchical nanoaggregate growth, we identify parameter regimes that favor rapid assembly of linear aggregates and map these regimes to sequence-defined candidate peptides for experimental synthesis and testing. This work establishes new understanding of assembly on previously unexplored time and length scales and presents an efficient and extensible protocol for computational screening and prediction of promising peptide chemistries to assemble nanostructures with desirable optoelectronic properties.

1. INTRODUCTION

Oligopeptides containing embedded π-conjugated subunits within the backbone have emerged as a valuable class of easily synthesized, biocompatible building blocks for triggerable self-assembly into micron-size β-sheet-like aggregates driven by hydrophobicity, π-stacking, and hydrogen bonding interactions[1–16]. Electronic delocalization between overlapping π-orbitals, proton transfer along hydrogen bonds[17], and potentially other more complex electronic effects[18] endow the supramolecular aggregates with emergent optical and electronic properties including fluorescence, electron transport, and exciton migration that make them attractive candidates for bioelectronic applications such as photovoltaics, energy harvesting and transport, biosensors, transistors, and light-emitting diodes[12,19–40]. The chemical diversity of peptide sequences and conjugated core chemistries – together with their water solubility, biocompatibility, and ease of synthesis – renders π-conjugated oligopeptides highly tunable building blocks that can be manipulated to control emergent structure and function[21,23,24,35,41–52].

Molecular interactions between the oligopeptide building blocks govern the assembly behavior into supramolecular aggregates of micron-size or larger on time scales of tens of seconds[41,53,54]. All-atom molecular simulations have proven extremely valuable in probing the molecular-level forces and mechanisms driving assembly such as the relative importance of aromatic and hydrogen bonding interactions in these and other systems of organic electronics[24,35,41,55–65]. Experimental work employing a range of techniques such as microrheology, circular dichroism, confocal fluorescence microscopy, cryo-electron microscopy, atomic force microscopy, laser scanning microscopy, transmission electron microscopy, spectrophotometry and fluorometry, has characterized macroscopic features of the supramolecular assemblies such as their critical fiber formation concentration, critical gel concentration, surface roughness, height profile, fibril thickness and aspect ratio, linear viscoelastic moduli, charge transport, hole mobility, electrical resistance, and absorption/emission spectra[14,24,32–34,39,41,54,66–70]. Bridging the disparate time and length scales between the micro-

scopic interactions probed by atomistic molecular dynamics and the macroscale structures characterized by experiment has proven challenging, but is of importance in determining how molecular chemistry influences and may be controlled to dictate supramolecular morphology and function.

One route to engage this problem is through coarse-grained molecular models that are more computationally efficient and permit access to far longer time and length scales than all-atom calculations. Since thermodynamic and folding behavior of many colloidal and protein systems depends only on a few key parameters[71–73], coarse-grained models can be designed to isolate, probe, and identify the most important effects governing self-assembly[74–77] and provide insight into the key microscopic interactions and physicochemical properties governing aggregation[78–80]. For example, using a bead-spring representation with additional angular flexibility constraints to model amyloid-type assembly, Ranganathan et al. showed that both interaction strength and stiffness are important factors governing the morphology of final self-assembled structures[78]. Sărić, Vácha, Frenkel and coworkers developed a series of top-down minimal models[81] to accurately predict the effect of different surface types on amyloid fibril formation[82], identify a proposed oligomer-dependent nucleation mechanism[83], probe the effect of surface coverage of fibrils on amyloid fibril nucleation and growth[80], and model pore formation in lipid membranes[84,85]. We have previously employed coarse-grained bead-level models of π-conjugated oligopeptides to reveal a hierarchical aggregation mechanism over length scales of tens of nanometers and time scales of hundreds of nanoseconds, fitted Smoluchowski coagulation models to extract aggregation rate constants and extrapolate the assembly dynamics, and studied the influence upon assembly of pH and flow[55,56].

Patchy particle models are a particular class of coarse-grained molecular models that have demonstrated great promise and flexibility in modeling colloids and peptides with directional interactions[81,86–94]. These models employ coarse-grained representations of molecular groups as typically spherical or ellipsoidal beads decorated by attractive surface patches that model directional–and potentially specific–intermolecular interactions. The effects of patch shape

and rotational entropy have been studied computationally and experimentally[95–97], as have the effects of decoupling rotational and translational diffusivity on pathways of aggregation[98]. Since they can reproduce experimental phase behavior, they have also been used to study phase transitions in colloids and proteins[99,100] and probe how the interactions and environmental conditions affect the phase diagrams of the resulting morphological structures[101–103]. Lomakin et al. introduced what they referred to as an "aelotopic" model of proteins, in which each protein was modeled as a sphere with interactions that depended on orientation, and showed that the directionality thus captured was responsible for generic features of protein aggregative properties in solution[104]. Hloucha et al. refined this template by introducing spheres with complementary potentials referred to as "patches-antipatches" to study virial coefficients in a system of bovine chymotrypsinogen in solution[105]. Kern and Frenkel[81] described a model of spheres decorated with directional square well patches to study the effects of directional interactions and interaction strength on the fluid–fluid coexistence curve of an abstract model system that could represent either colloids or proteins. In a similar vein, Zhang and Glotzer showed that using small rigid spheres as the source of directional interactions could be tuned to result in a plethora of different self-assembled structures[86]. Using a patch-antipatch idea, Dorsaz et al.[92] successfully modeled particular aspects of protein crystallization, and Morgan et al.[106] derived a minimalist design rule for the assembly of a Bernal spiral. Long et al. used a machine learning approach to design patchy colloids to assemble into specific target structures[107–110]. Xu et al. and Liu et al. probed the dependence of self-assembled structures on aspect ratios of the central body and produced phase diagrams with respect to aspect ratios and concentrations[111,112], while Carpency, Gunton, and Rickman studied the separate effects of shape anisotropy and patchiness on the phase transitions of two-patch and four-patch Kern-Frenkel patchy particles[113].

In this work, we develop a patchy particle model with which to study the hierarchical self-assembly of the DXXX-Π-XXXD family of π-conjugated oligopeptides into β-sheet-like aggregates with optical and electronic functionality[24,35,41,42,53–56,70] (Fig. 1a). The Π subunit

represents a π-conjugated core such as oligophenylenevinylene (OPV) or perylenediimide (PDI) and the X represents one of the 20 natural amino acids. We further restrict our study to symmetric chemistries D-X_3-X_2-X_1-Π-X_1-X_2-X_3-D where the N-to-C directionality of the tetrapeptide wings proceeds away from the core such that the molecule possesses two Asp residue C-termini. The number of peptides in the family is $(p \times 20^3)$, where p is the number of Π cores to be considered. At high pH (pH $\gtrsim$ 7) the C-termini and terminal Asp side chain carboxyl groups are deprotonated and each peptide carries a (-4) charge that prohibits large-scale assembly by Coulombic repulsion; these groups protonate at low pH (pH $\lesssim$ 1) to produce neutral oligopeptides and trigger supramolecular assembly by hydrophobicity, π-stacking, and hydrogen bonding[24,35,41,53]. The morphology and optoelectronic functionality of the terminal supramolecular aggregates can be controlled by the chemistry of the π-conjugated core and amino acid sequence of the peptide wings[24,41,54,70]. Prior computational studies have employed all-atom and united atom (i.e., bead-level) molecular models to study the morphology, thermodynamics, and kinetics of assembly[24,35,41,42,55,56], but the computational expense associated with these calculations limited study to a restricted number of chemistries in simulations comprising a few hundred monomers over time scales of 100 ns and length scales of 10 nm. In contrast, our patchy particle model is capable of reaching time scales of 100 μs and length scales of 100 nm in simulations comprising 10,000 monomers. The model is also sufficiently computationally inexpensive to efficiently scan over a broad parameter range and screen the diversity of molecular chemistries within the DXXX-Π-XXXD family. The patchy particle model therefore exists at a resolution that is inexpensive enough to efficiently scan through chemical space and probe mesoscopic length and time scales, while retaining sufficient molecular detail to allow us to resolve the microscopic physicochemical determinants of assembly behavior. We exploit these properties of the model to pursue the main goals of this work: (i) to identify parameter regimes of our patchy particle model that promote favorable self-assembly, and (ii) to translate these results into new molecular understanding of assembly and design precepts for rational engineering

of novel oligopeptides.

The outline of this paper is as follows. In Section 2, we describe the construction of the patchy particle model, detail our molecular simulation methods, and provide structural definitions for the aggregation hierarchy. In Section 3, we present the results of our virtual high-throughput screen through the model parameter space to identify parameter regimes that lead to rapid assembly of linear β-sheet-like aggregates, and discuss the mapping of these parameters to particular oligopeptide chemistries. In Section 4, we present our conclusions and directions for future work.

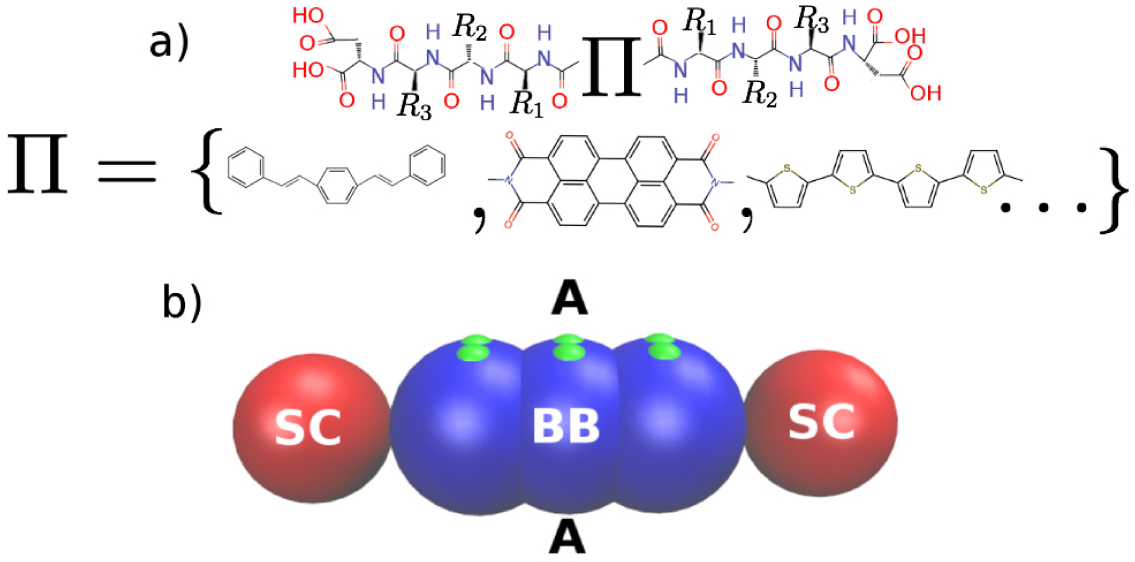


Figure 1: Chemical structure and patchy particle model of the DXXX-Π-XXXD π-conjugated oligopeptides. (a) Chemical structure of a generic DXXX-Π-XXXD oligopeptide. The peptide wings are mirror symmetric in that the N-to-C directionality progresses away from the π-conjugated core and the tetrapeptide sequences are identical. The side chains $\{R_1, R_2, R_3\}$ may be tuned to any of the 20 naturally occurring amino acids and Π is a π-conjugated core such as oligophenylenevinylene (OPV), perylenediimide (PDI), or quaterthiophene (OT4). The four terminal carboxyl groups belonging to the two C-termini and distal Asp residue side chains deprotonate at high pH (pH $\gtrsim$ 7) and protonate at low pH (pH $\lesssim$ 1) enabling pH-triggerable assembly. Image produced using Marvin 18.10.0 (ChemAxon, 2018) (`http://www.chemaxon.com`). (b) Patchy particle model of a generic DXXX-Π-XXXD oligopeptide. The small green *A* beads located on the upper and lower face of the model model cofacial aromatic interactions, the large red *SC* beads model interactions of the peptide wings, and the large blue *BB* beads model non-cofacial aromatic core interactions. The diameter of the *SC* beads is a parameter of the model. The length of the core is twice the diameter of a single *BB* bead defined as the minimum of the Lennard-Jones *BB* interaction potential. Image produced using VMD[114].

2. METHODS

2.1 Patchy particle model of DXXX-Π-XXXD oligopeptides

We formulated our patchy particle model with three major design principles in mind. First, so as to enable inexpensive large-scale molecular simulation and limit the extent of parameter screening required for model parameterization, we assert that the model contain a relatively small number of parameters controlling the salient features of DXXX-Π-XXXD oligopeptide assembly: the linear geometry of the molecule, the directional character of cofacial core–core interactions, and the variability of peptide side chain chemistries. Second, the model should reproduce the gross structural features of aggregation observed in prior all-atom and coarse-grained simulations[35,55,56]. In particular, the observation of branched fibrillar structures with a fractal dimension of D ∼1.5 on length scales of ∼10 nm and a hierarchy of the growth of different cluster types. Third, while the model may not maintain amino acid-level resolution, it should admit an approximate mapping from our previously-established Martini model for DXXX-Π-XXXD peptides that does maintain distinct amino acid identities[55]. The difference in resolution between the two models means that the mapping from the higher-resolution Martini model to the lower-resolution patchy particle model may be non-injective (i.e., not one-to-one, with the result that multiple Martini peptides correspond to the same patchy particle model) and non-surjective (i.e., particular parameterizations of the patchy particle model may not correspond to the Martini model of any realizable oligopeptide). Nevertheless, the existence of an approximate mapping is vital in both defining the physically realizable parameter range in the patchy particle model (i.e., forward mapping), and in identifying realizable oligopeptide chemistries corresponding to promising parameter regimes of the patchy particle model (i.e., reverse mapping).

2.1.1. Model construction

Building upon previous patchy particle models of self-assembling peptides established by Vácha, Sărić, and Frenkel[80,83,115] and of self-assembling colloids established by Zhang and Glotzer[86], we construct a patchy particle model of a DXXX-Π-XXXD oligopeptide as illustrated in Fig. 1b. Three overlapping spherical backbone (BB) beads represent the three (poly)aromatic rings comprising a π-conjugated core. Six sticky patches represented by rigidly attached aromatic (A) beads are placed on the top and bottom of the long axis of the three BB beads to represent the directionally-specific cofacial π-π interactions of the aromatic rings. In principle, the arrangements of the spheres comprising the sticky patches could be modified to probe the effects of aromatic core geometry, similar to Sărić et al.[80], but we do not study this effect here. A single large side chain (SC) bead affixed rigidly on either end of the three core beads represents the peptide wings. Since charges on the symmetric peptide wings are known to prevent large-scale assembly due to unfavorable Coulombic interactions[24,42], for the purposes of this work we consider only uncharged SC beads and model interactions between the three bead types using 12-6 Lennard-Jones (LJ) potentials[116].

It is clear that the highly coarse-grained nature of this model integrates out many of the atomistic level details of the peptides. Most egregiously, the patchy particle model is fully rigid and the peptide wings represented by a single sphere. Nevertheless, the model is designed to retain the essential molecular geometry – an elongated monomer comprising peptidic and aromatic subunits – and interactions – peptide wings capable of mutual stacking and a π-conjugated core with preferred parallel stacking orientations. We demonstrate in Section 3.2.3 that despite its simplicity, this relatively simplistic model is capable of reproducing the salient features of peptide aggregation. The trade-off for this simplicity is access to vastly longer length and time scales than would be possible with higher resolution descriptions. In particular, our patchy particle model permitted simulation of about ten thousand monomers on length scales of hundreds of nanometers and time scales of hundreds of microseconds. This represents approximately 250× longer length scales and 2000× longer

time scales than were accessible in previous studies[35,55,56], enabling access to long-time and many-body behaviors out of reach to finer resolution models. Furthermore, the simplicity of the model is a virtue in interpretably illuminating the effect of different model parameters on self-assembled morphologies and growth kinetics. Finally, the inexpensive model can be used to rapidly and efficiently explore different parameter regimes to search for combinations producing desirable assembly behaviors, and then the parameters mapped back to physically realizable peptide chemistries for investigation by higher resolution calculations and ultimately experimentation (see Section 3.3).

We now describe the fixed parameters of the model, which are set by appealing to simple geometric considerations and by analogy with the DFAG-OPV3-GAFD system studied previously[35,55,56], which consists of mirror symmetric peptide wings comprising Ala, Gly, and Phe residues flanking an oligo(p-phenylenevinylene)3 (OPV3) aromatic core. The mass of a single BB bead is set to $m^* = 108$ amu, close to one-third the mass of an OPV3 core (~ 112 amu), while the mass of a single SC bead is set to $3.75m^* = 405$ amu, close to the mass of a DFAG tetrapeptide (~ 407 amu). The A beads carry no mass and are virtual attractive particles. The LJ diameter of the BB beads is set to $\sigma_{BB} = 2^{-1/6}$ nm, such that the length of the aromatic core $d_{e\to e} = 2d_{\min}^{\mathrm{LJ}} = 2\left(2^{1/6}\sigma_{BB}^{\mathrm{LJ}}\right) = 2 \times (2^{1/6} \times 2^{-1/6}\ \mathrm{nm}) = 2$ nm is approximately that of an OPV3 core (~ 1.88 nm)[55]. The LJ diameter of the A beads is set to $\sigma_A = 2^{5/6}/8$ nm and are centered at a radial distance of 0.475 nm from the center of the BB beads in order to mitigate their protusion above the surface of the BB beads and therefore any spurious geometric effects associated with corrugations of the BB bead surface and interlocking of the patchy particles[117]. The cutoff radius for LJ interactions is set to 2.0 nm. The beads constituting the peptide define a rigid body.

2.1.2. Model parameterization

The remaining free parameters in the model are the LJ diameter of the side chain beads, σ_{SC}, the LJ well depth of the beads representing the side chains, ϵ_{SC}, the LJ well depth of

the beads representing non-cofacial aromatic core interactions, ϵ_{BB}, and the LJ well depth of beads representing cofacial aromatic core interactions, ϵ_A (Fig. 1b). We describe in this section how we defined the range of appropriate values for these parameters in order to span the chemical space of realizable π-conjugated cores and peptide wings.

Side chain diameter, σ_{SC}.

This parameter represents the van der Waals volume occupied by the tetrapeptide wing, encapsulated as the size of the SC sphere. We determine an appropriate range of values for this parameter by estimating the solvent-accessible surface area (SASA) for amino acid residues modeled by the coarse-grained Martini force field (Section 2.2). We specify σ_{SC} as the diameter of a sphere possessing the same surface area as that computed for the Martini model of the tetrapeptide. We verified in the case of DFAG that for such short peptides the measured SASA of the complete tetrapeptide chain is equal within error bars to the sum of the SASAs computed for the individual amino acid residues. Employing this procedure, we identified σ_{DAAA} = 1.25 nm for the smallest A residue in the Martini model, and σ_{DWWW} = 1.63 nm for the largest W residue. Accordingly, we identify σ_{SC} = 1.0-1.75 nm as an appropriate physically meaningful range over which to to study the excluded volume of the peptide wings.

Side chain well depth, ϵ_{SC}.

This parameter characterizes the interaction strength between peptide wings. An appropriate range for this parameter may be estimated by constructing predictions for the dimerization potential of mean force (PMF) of isolated DXXX tetrapeptides. These calculations also provide a mapping from a sequence-defined chemistry to a patchy particle model, and can be used to perform the reverse mapping to identify candidate chemistries corresponding to particularly promising patchy particle parameterizations (see Section 3.3). We define an appropriate range of ϵ_{SC} by appealing to our prior work in which we trained a quantita-

tive structure property relationship (QSPR) model to accurately predict the dimerization free energies for nonpolar DXXX-NDI-XXXD and DXXX-PDI-XXXD oligopeptides[42]. The model was trained over dimerization PMFs computed for a subset of peptide sequences using molecular dynamics simulations of all-atom oligopeptide representations in implicit solvent. To extract from the QSPR model an estimate of the dimerization free energy of the peptide wings alone ΔF_{wing}, we compute dimerization PMFs for isolated NDI and PDI cores using the simulation protocol detailed in Ref.[42] and subtract these values of ΔF_{core} from the QSPR model predictions. For a particular tetrapeptide sequence, we take the average of the predictions from the NDI and PDI QSPR models and divide the result in half to provide an estimate of ϵ_{SC} for a single DXXX peptide wing. The dimerization PMF is found to span a range of ΔF_{wing} = (-60)-5 k_BT at T = 298 K over all nonpolar peptide sequences considered. Since repulsive interactions and overly attractive interactions are both known to lead to poorly aligned oligopeptide aggregates[42,56], this motivates us to consider a parameter range of ϵ_{SC} = 0.2-10 k_BT at T = 298 K.

Non-cofacial aromatic interaction well depth, ϵ_{BB}.

This parameter controls the non-cofacial (i.e., non-parallel stacked) interaction strength between the backbone particles constituting the π-conjugated aromatic core. Umbrella sampling calculations conducted using a Martini model of an OPV3 core reveal a $\Delta F_{\text{core}}^{\text{ncf}}$ = (-9)-(-3) k_BT free energy well for non-cofacial dimerization at T = 298 K (see Section 2.2). Since our patchy particle model comprises three BB beads, this suggests that we employ ϵ_{BB} = 1-3 k_BT. This is in line with the Lennard-Jones energy parameter of ϵ = 1 k_BT corresponding to the Martini SC5 bead used to represent the two C atoms comprising one edge of an aromatic ring[118,119]. Although an attractive ϵ_{BB} is found to be essential in promoting the formation of well-aligned aggregates, computational exploration also reveals the dependence of cluster growth rates and morphologies to be quite insensitive to its value – a 6000% change in ϵ_{BB} over the range 0.01–0.6 k_BT produced only a ∼16% change in cluster

growth rates – motivating us to fix it at $\epsilon_{BB} = 1\ k_BT$ at $T = 298$ K.

Cofacial aromatic interaction well depth, ϵ_A.

This parameter controls the cofacial (i.e., parallel stacked) interaction strength between the backbone particles constituting the π-conjugated aromatic core. Umbrella sampling calculations conducted using a Martini model of an OPV3 core reveal a $\Delta F^{\mathrm{cf}}_{\mathrm{core}}$ = (-18) k_BT free energy well for cofacial dimerization at $T = 298$ K (Section 2.2). Eliminating the (-3) k_BT attributable to the non-cofacial aromatic interactions, this suggests a cofacial interaction free energy of $\sim$2.5 k_BT for each of the two A beads decorating each side of each of the three BB particles. The strength of this interaction can be modulated by changing the chemistry or number of fused aromatic rings in each subunit of the core. Accordingly, we elect to consider a 300% change in this parameter to scan over ϵ_A = 2.5-7.5 k_BT at T = 298 K in order to study the influence of the strength of π-π stacking interactions upon assembly.

2.2 Martini model simulations in Gromacs

Coarse-grained simulations of the oligopeptides were conducted using the Martini model in which approximately four atoms are lumped to each coarse-grained bead[120]. This model resolution presents a judicious balance between molecular realism and efficient simulation. We have previously constructed a model of the DFAG-OPV3-GAFD chemistry based on the Martini model that we explicitly reparameterized against all-atom simulations to better reproduce the molecular level thermodynamics[55]. We also conducted simulations of single amino acids represented in the original Martini model version 2.2[119] to help inform parameter selection. Calculations were performed in the Gromacs 4.6 suite[121], employing Martini polarizable water[122]. Simulations were conducted in an NPT ensemble at 298 K and 1 atm, employing a velocity rescaling thermostat[123] and a Parrinello-Rahman barostat[124]. Three dimensional periodic boundary conditions were employed. Electrostatics were treated with the

reaction field method with $\epsilon_r = 2.5$ and $\epsilon_{rf} = \infty$[122]. Lennard-Jones interactions were shifted smoothly to zero at $r_{VDW} = 1.1$ nm. The classical equations of motion were integrated using the leap-frog algorithm[125] with a time step of 5 fs.

Solvent accessible surface area calculations used to estimate σ_{SC} were conducted in $5.7 \times 5.7 \times 5.7$ nm^3 boxes using the Gromacs gmx SASA tool[126,127] with the radii of the beads set to one half of their Martini LJ σ parameters. Estimates for ϵ_{BB} and ϵ_A were generated from umbrella sampling calculations in $10 \times 10 \times 16$ nm^3 boxes to compute the dimerization PMF well depth for cofacial (i.e., parallel stacked) and non-cofacial (i.e., edge stacked) dimerization of the OPV3 cores. Umbrella sampling along the dimerization pathway was conducted in windows at 0.1 nm increments in the center of mass separation between the peptides along the z-direction, employing harmonic restraints of $k_{\text{umb}} = 10^3$ kJ/mol.nm^2 parallel to the pulling direction and $k_{\text{rest}} = 10^4$ kJ/mol.nm^2 perpendicular to the pulling direction in order to maintain in-register stacking. A 1 ns equilibration run was conducted in each umbrella window followed by a 14 ns production run. The PMF was estimated to within a tolerance of 10^{-6} by combining the umbrella sampling data using the Weighted Histogram Analysis Method (WHAM)[128]. Uncertainties were estimated by block averaging. Simulations were conducted on a single Intel Xeon E5-2660 2.2 GHz core achieving execution rates of 17.5 ns/day.

2.3 Patchy particle simulations in HOOMD

Simulations of peptides modeled by the patchy particle model were conducted in HOOMD 2.1.7[129,130]. We employ reduced units where the unit of distance is $d^* = 1$ nm, the unit of mass is $m^* = 108$ amu, and the unit of energy is such that $\epsilon^*/k_BT = 1.0$ at 298 K, from which we compute a reduced unit of time $\tau^* = \sqrt{\frac{m^*(d^*)^2}{\epsilon^*}} = 6.6$ ps. The coarse-grained nature of the patchy particle model integrates out configurational degrees of freedom and also smooths the underlying potential energy landscape, leading to artificial acceleration of the system dynamics[120,131,132]. We have previously shown in simulations of DFAG-OPV3-GAFD peptides that

there is no significant speedup in moving from an all-atom to a Martini model description[55]. We use a similar approach here to ascertain the speedup in moving from the Martini model to a patchy particle description. Specifically, we conduct simulations of isolated peptides at 298 K and 1 bar under each model, track the mean squared displacement, and employ the Einstein relation[116] to estimate translational diffusion coefficients of $D_{\text{Martini}} = (7 \pm 2) \times 10^{-6}$ cm^2/s and $D_{\text{patchy}} = (3.4 \pm 0.9) \times 10^{-2}$ cm^2/s. Matching the diffusivities implies a $\sim$5000$\times$ speedup of the patchy particle model relative to Martini, and therefore all-atom, time scales. The translational self-diffusion is likely to be the most salient quality determining speedup since it governs the mean free path between the collisions necessary for aggregation. Accordingly, each HOOMD time step of $\tau^* = 6.6$ ps corresponds to $\tau^*_{\text{patchy}} \approx 33$ ns. In the remainder of this article, we correct for this speedup when reporting time in real units. We note that this speedup was a primary motivating factor for the development of the patchy particle model, as it enables access to orders of magnitude longer time and length scales than those attainable by higher resolution models.

It is a primary goal of this study to explore how values of ϵ_A, ϵ_{SC}, and σ_{SC}, corresponding to different oligopeptide chemistries, influence the morphology and kinetics of peptide aggregation. To this end, we sweep over the physically motivated ranges of each of these by conducting five independent simulations at each of the $5 \times 4 \times 3 = 60$ parameter combinations listed in Table 1. Simulations are initialized by arranging 10,648 patchy particles over a uniform cubic lattice within a $158 \times 158 \times 158$ nm^3 cubic box, corresponding to a concentration of 4.4 mM. Experimental studies of oligopeptide assembly have been conducted up to 0.86 mM[54]. We study five-fold higher concentrations in order to accelerate and better observe large-scale assembly within our simulation cell. Further, we have previously demonstrated that the aggregation mechanism of DFAG-OPV3-GAFD peptides is independent of concentration over the range 5–44 mM[55]. Langevin dynamics simulations are conducted from the initial monodisperse state at a reduced temperature of $T^* = 1$, with diameter-scaled damping coefficients of $\gamma_i = \lambda d_i$, where $\lambda = 1\ m^*/d^*\tau^*$, and d_i is the LJ minimum of the ith particle.

A total of 2×10^7 Langevin steps of $dt^* = 10^{-3}$ are performed, corresponding to 660 μs of simulation. After the first 6×10^5 time steps (19.8 μs) are discarded for equilibration, after which time the temperature and pressure attain stable values and the remaining 1.94×10^7 time steps (640.2 μs) are allocated to production runs. Simulations were conducted on 4 × NVIDIA GK110 (K20X) "Kepler" GPUs on the Blue Waters supercomputer at the University of Illinois at Urbana-Champaign achieving execution rates of 240 time steps per second (∼680 μs/day).

Table 1: Table of Patchy Parameters[1]

ϵ_{SC} (k_BT)	σ_{SC} (nm)	ϵ_A (k_BT)
0.2	1.00	2.5
0.9	1.25	5.0
2.0	1.50	7.5
6.0	1.75	
10.0		

2.4 Cluster types

We have previously defined hierarchical criteria by which to judge whether pairs of patchy particles should be judged to be associated into a single aggregate. In particular, we defined three different classes of cluster: (i) aligned clusters, which are made up of molecules with well-aligned aromatic cores, (ii) optical clusters, which are made of molecules with proximate but not necessarily well-stacked aromatic cores, and (iii) contact clusters, which are made up of molecules with any proximate beads in the aromatic core or peptide wings[55,56]. Given the coarse resolution of our patchy particle model, we focus in this work on two cluster definitions: contact clusters and optical clusters.

Contact clusters. Two peptides are defined to belong to the same contact cluster if the intermolecular distance between any pair of beads is below a cutoff distance threshold. This

[1]Parameter values employed in the patchy particle model parameter sweep. A total of $5 \times 4 \times 3 = 60$ different parameter sets were considered corresponding to all ϵ_A, ϵ_{SC}, and σ_{SC} combinations. Five independent simulations were performed at each parameter set.

presents a relatively loose definition of association, since the peptides cores are not necessarily mutually well aligned and therefore may lack the π-π stacking and electronic delocalization necessary to endow the aggregate with optoelectronic functionality. Mathematically, a pair of peptides a and b reside in a contact cluster if $R_{a,b}^{\text{con}} < R_{\text{cut}}^{\text{con}}$, where,

$$R_{a,b}^{\text{con}} = \min_{i \in a} \min_{j \in b} r_{ij}, \tag{1}$$

wherein r_{ij} is the distance between beads i and j. We specify $R_{\text{cut}}^{\text{con}} = \max(2^{1/6}\sigma_{BB} + 0.1$ nm, $2^{1/6}\sigma_{SC} + 0.1$ nm) such that two monomers are considered to be in a cluster if two of their beads are within 0.1 nm of the LJ minima of the larger of the SC and BB beads.

Optical clusters. Two peptides are defined to belong to the same optical cluster if the distance between any pair of A beads is below a cutoff distance threshold. This metric assures that the peptide cores are in close proximity and interacting approximately cofacially and may therefore exhibit optoelectronic functionality, although they do not necessarily display perfect in-register stacking. Specifically, a pair of peptides a and b are in an optical cluster if $R_{a,b}^{\text{opt}} < R_{\text{cut}}^{\text{opt}}$, where,

$$R_{a,b}^{\text{opt}} = \min_{i \in (A \text{ beads} \in a)} \min_{j \in (A \text{ beads} \in b)} r_{ij}, \tag{2}$$

and $R_{\text{cut}}^{\text{opt}} = 0.35$ nm. This somewhat restrictive cutoff is chosen because the A beads are centered at 0.475 nm from the center of the BB beads, so two monomers reside within the same optical cluster if the BB bead centers lie within $(0.35 + 2{\times}0.475) = 1.3$ nm. This metric therefore embodies both core proximity and cofacial interaction. By construction, the metric defining an optical cluster nests it within a contact cluster: peptides in an optical cluster are also in a contact cluster, but the inverse is not true. Accordingly, optical clusters are constrained to be equal or smaller in size to the contact cluster within which they reside.

3. RESULTS AND DISCUSSION

The primary goals of this work are to (i) discern what parameter regimes of our patchy particle model promote favorable self-assembly rates and morphologies, and (ii) translate these findings into new understanding of large-scale and long-time assembly and design precepts for rational design of candidate oligopeptide chemistries. We engage these goals by determining how the side chain well depth ϵ_{SC}, side chain diameter σ_{SC}, and cofacial aromatic well depth ϵ_A affect the mechanisms, kinetics, and morphology of self-assembly in our patchy particle model of DXXX-Π-XXXD oligopeptides. We focus on the growth rate of optical and contact clusters and the fractal nature of the aggregates formed. We use these findings to draw conclusions about the physical principles governing aggregation and how to modify the interactions through judicious selection of peptide chemistry to promote desirable assembly behaviors.

We focus on the kinetics and morphology of assembly rather than the thermodynamics because the emergent large-scale aggregates are expected to be out-of-equilibrium, kinetically-trapped states rather than globally-stable thermodynamic minima. For the majority of supramolecular self-assembled aggregates formed from small molecule building blocks, the thermodynamic minima is expected to be a crystal whereas the commonly observed self-assembled aggregates are frequently metastable, kinetically-trapped gel-like or fibrillar states[133–135]. The latter morphologies are often of engineering interest since the pathway dependence of their formation can be exploited to produce tunable and responsive aggregates[136,137], and their formation is consistent with a model in which assembly is under kinetic control[56,70,134,135]. The pathway dependence of assembly has been previously exploited in the case of the DXXX-Π-XXXD family wherein the final aggregate morphology has been tuned by the presence or absence of an external flow field to modulates the extent of branched versus linear aggregates[54]. We have previously conducted thermodynamic stability analyses of small oligomers of up to five peptides whose structure does approximate equilibrated as-

semblies[35,42]. Although interesting in quantifying (meta)stability, a similar analysis of the large scale kinetically-trapped aggregates is less illuminating than a kinetic approach that seeks to quantify the rate and morphology of assembly without appealing to thermodynamic stability concerns.

3.1 Absolute and relative growth rate of optical clusters

The supramolecular aggregates formed by peptide self-assembly are endowed with opto-electronic functionality by electronic delocalization over the parallel stacked π-conjugated cores[19–22]. Our definition of optical clusters identifies such well-stacked configurations, distinct from the less well-aligned aggregates identified merely as contact clusters (Section 2.4). In order to promote the rapid formation of well-aligned supramolecular aggregates, we adopt as our dual objective functions (i) the absolute growth rate of optical clusters, and (ii) the growth rate of optical clusters relative to contact clusters. Quantification and maximization of these two measures with respect to the patchy particle model parameters ϵ_{SC}, σ_{SC}, and ϵ_A allows us to identify parameter regimes in which we observe rapid formation of well-aligned optical clusters, but not at the expense of poorly aligned aggregation into disordered contact clusters. It is desirable that the growth rate of optical clusters be both fast and as close as possible to that of contact clusters since this corresponds to rapid growth of aggregates with well-aligned interacting aromatic cores. We recall that the two cluster measures are hierarchical: optical clusters are also contact clusters, but the inverse is not true. Accordingly, optical cluster size and growth rate is constrained to be smaller than or equal to contact cluster size and growth rate.

3.1.1. Modeling growth kinetics by Smoluchowski coagulation theory

We quantify the formation rates of optical and contact clusters by tracking the mass-averaged cluster size[56,138,139],

$$\mu_2(t) = \frac{1}{m_{\text{mon}}} \frac{\sum_{i=1}^{N} m_i^2 N_i(t)}{\sum_{i=1}^{N} m_i N_i(t)} = \frac{\sum_{i=1}^{N} i^2 N_i(t)}{\sum_{i=1}^{N} i N_i(t)}, \tag{3}$$

where $m_i = m_{\text{mon}} i$ is the mass of a cluster containing i monomers of mass m_{mon}, $N_i(t)$ is the number of clusters of size i in the simulation at time t, and cluster size is reported in terms of the number of monomers. The mass-averaged cluster size is an experimentally-accessible observable that is often preferred over the number-average μ_1 as it is less sensitive to fluctuations in the numbers of light clusters[140]. For each of our 60 parameter settings (Table 1), we track the time evolution of $\mu_2^{\text{opt}}(t)$ and $\mu_2^{\text{con}}(t)$. Fig. 2 presents illustrative examples of the optical and contact cluster time evolution for four selected parameter sets..

We have previously employed the Smoluchowski coagulation theory of irreversible aggregation to model cluster formation[138,141–148] in smaller systems comprising hundreds of monomers over hundreds of nanoseconds[55,56]. We note that the good description afforded by a model that does not include fragmentation supports a model of kinetically controlled aggregation leading to the formation of kinetically-trapped aggregates. The continuous time Smoluchowski coagulation equation for irreversible aggregation and discrete aggregate sizes is[55,56,141],

$$\frac{dn_r}{dt} = \frac{1}{2} \sum_{i=1}^{r-1} K_{i,r-i} n_i(t) n_{r-i}(t) - \sum_{i=1}^{\infty} K_{r,i} n_r(t) n_i(t), \tag{4}$$

where t is time, $K_{i,j}$ is a second order rate constant kernel for the association of two aggregates of size i and j, and $n_r(t)$ is the number concentration of aggregates of size r at time t. The observed trends in μ_2 indicate that the Smoluchowski model also presents an appropriate description for the formation rate of optical – after a $\sim$250 μs transient – and contact clusters in these much larger systems of nearly 11,000 monomers over hundreds of microseconds. In

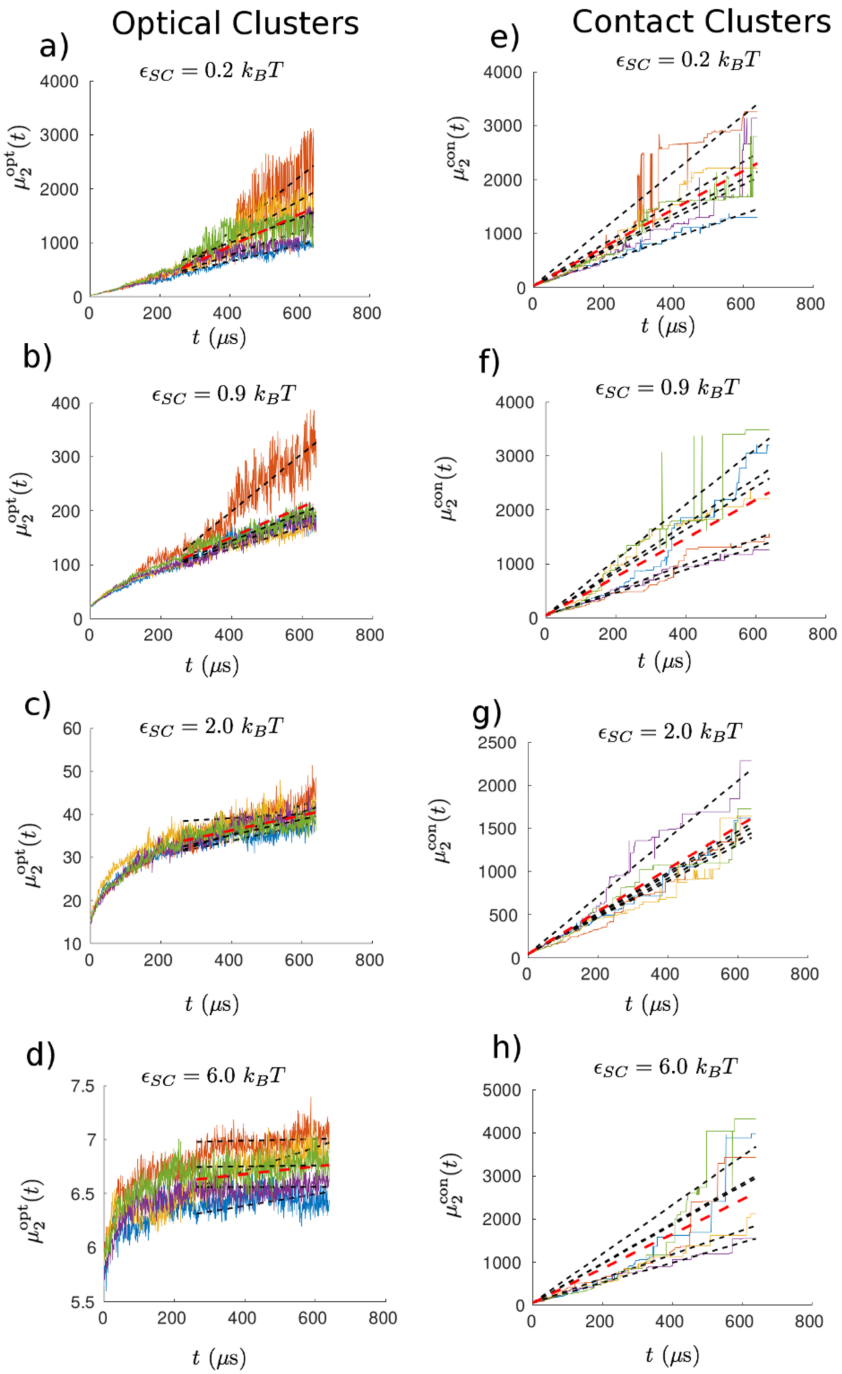


Figure 2: Time evolution of the mass-averaged cluster size μ_2 for (a)-(d) optical clusters and (e)-(h) contact clusters at parameters $\epsilon_A = 2.5\ k_BT$, $\sigma_{SC} = 1.5$ nm and four different possible values of ϵ_{SC} noted above each panel. Each of the five independent runs under each condition is shown in a different color. Black dashed lines correspond to the separate fits to the independent runs, of the form $\mu_2(t) = \mu_2(t_0) + k_{\mathrm{opt}}(t - t_0)$. Thick red dashed lines correspond to the average fit. t_0 is specified as 264 μs for optical clusters and 0 μs for contact clusters.

particular, the linear growth kinetics are consistent with a size-independent kernel $K_{i,j} = K$, for which the analytical prediction of the Smoluchowski model for the mass-averaged cluster size under arbitrary initial conditions is[138,141,143],

$$\mu_2(t) = \mu_2(t_0) + k(t - t_0), \tag{5}$$

where $\mu_2(t_0)$ is the mass-averaged cluster size at t_0, $k = KM_1 = 2/t_c$ is the lumped characteristic coagulation rate constant incorporating the effect of concentration[146], M_1 is the total concentration of monomers in the system, and t_c is the characteristic coagulation time[138]. We superpose the best least squares fit of Eqn. 5 to the recorded mass averaged optical and contact cluster size for the four selected parameter sets presented in Fig. 2 to illustrate empirical estimation of k and t_0 from our data.

We present in Figs. 3 and 4 the best fit values of k_{opt} and k_{con} at each of the 60 parameter settings, and in Fig. 5 the ratio of these rates. We recall that our dual objectives are maximization of k_{opt} (Fig. 3) and $k_{\text{opt}}/k_{\text{con}}$ (Fig. 5). In performing the fits, we adopt $t_0^{\text{con}} = 0$ μs and $t_0^{\text{opt}} = 264$ μs to reflect the presence of the initial transient in optical cluster formation. The origin of this transient is the different length scales of aggregation (see also Section 3.2) and the different ways in which optical clusters can form, either by initial collision of two clusters in a manner that allows the A beads to immediately interact, or by rearrangement of monomers within contact clusters to accommodate these cofacial interactions. There is an initially rapid increase in optical cluster size due to rapid agglomeration of peptide monomers from the monodisperse state into small contact clusters in which mutual core alignment of monomers among small aggregates is relatively fast, and formation of optical clusters proceeds by both collision and monomer rearrangement. This is succeeded by a slower growth phase in which aggregation proceeds by the collision of larger aggregates. The morphology of the small-scale aggregates controls the likelihood that there will be exposed A beads, which in turns controls the likelihood that two colliding aggregates will interact via

cofacial core–core interactions and form an optical cluster. Time scales for rearrangement of monomers in these larger clusters is much longer than that for rearrangement in the smaller clusters, so the slowdown in the growth rate occurs due to the fact that after a certain size, optical clusters tend to form primarily through collision. This effect is responsible for the slowdown at longer times of the initially rapid optical cluster formation rate. We previously noted a similar effect that we ascribed to the decrease in optical cluster mobility and available surface area with increasing size and charge, and which we modeled using a size-dependent optical cluster growth rate[55,56]. We were unable to reach long enough time and length scales in our prior work to be able to decisively distinguish between a discrete change in growth mechanism and a continuous one; here on time scales of hundreds of microseconds we definitively show a clear convergence to linear growth after an initial transient. Contact cluster formation, on the other hand, exhibits no initial transient since aggregation is largely independent of the microscopic details of the cluster morphology, with cluster formation simply requiring that any two beads in the clusters lie within the cutoff distance threshold.

3.1.2. Small ϵ_{SC} promotes rapid optical cluster growth.

Inspection of Fig. 3 illustrates that the side chain interaction strength ϵ_{SC} has the strongest influence on the optical cluster growth rate. At fixed σ_{SC} and ϵ_A, reducing ϵ_{SC} to make the side chains less strongly interacting results in an increase of k_{opt} and faster formation of optical clusters. The increase in growth rate goes up significantly once ϵ_{SC} is lowered beneath $\epsilon_{BB} = 1\ k_BT$. The most pronounced increase in assembly rate occurs upon lowering ϵ_{SC} from 0.9 k_BT to 0.2 k_BT, which results in an average 30-fold increase of k_{opt} per k_BT, compared to an average 5-fold increase per k_BT for lowering ϵ_{SC} from 2 k_BT to 0.9 k_BT and from 6 k_BT to 2 k_BT, and only an average 1.2-fold increase per k_BT for lowering ϵ_{SC} from 10 k_BT to 6 k_BT. Recalling that the interaction strength of the core (BB) beads is $\epsilon_{BB} = 1$ k_BT, this observation is consistent with a mechanism in which reducing the "stickiness" of the side chains below that of the cores begins to promote well-aligned cluster assembly through

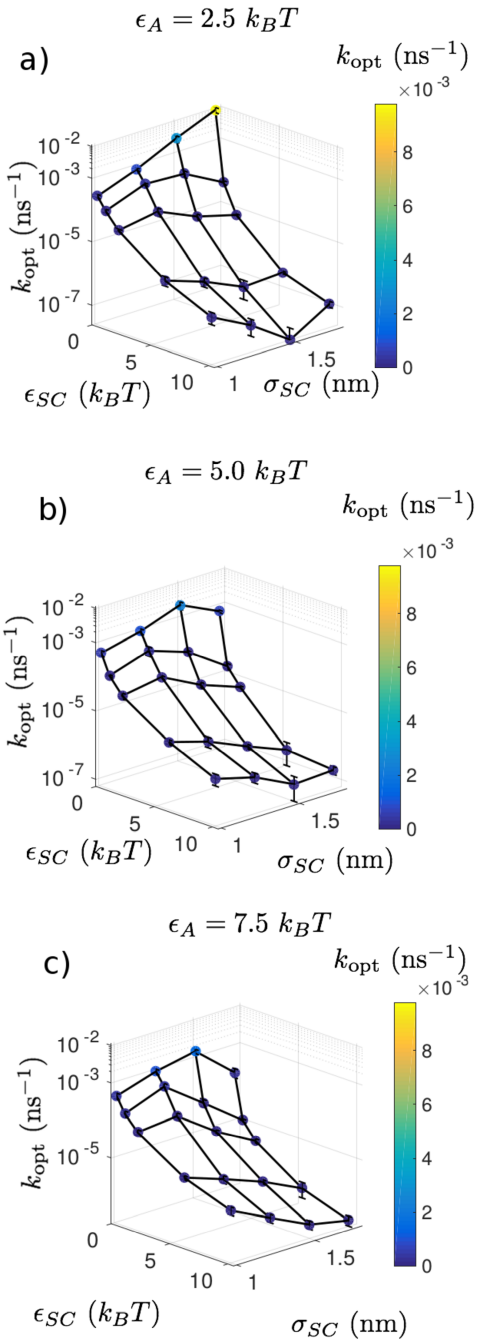


Figure 3: Best fit rate constants k_{opt} for the formation of optical clusters as a function of patchy particle model parameters ϵ_{SC} and σ_{SC} at (a) $\epsilon_A = 2.5\ k_BT$, (b) $\epsilon_A = 5.0\ k_BT$, and (c) $\epsilon_A = 7.5\ k_BT$. Mean values of k_{opt} are calculated over five independent simulations at each parameter setting and plotted on logarithmic axes to better illuminate the trends. Relative standard errors in k_{opt} over the five runs are estimated as 6-133%.

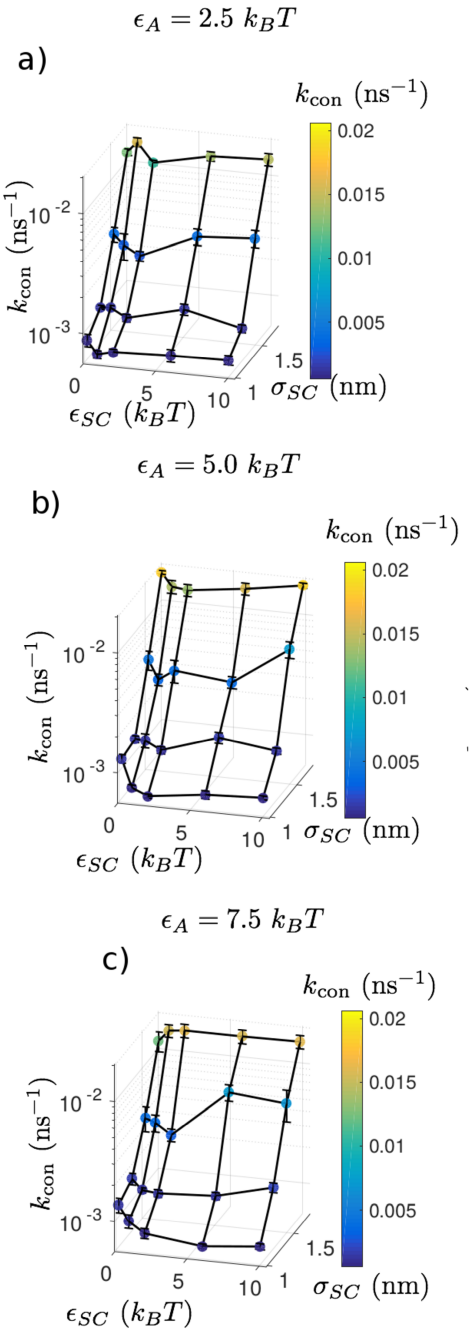


Figure 4: Best fit rate constants k_{con} for the formation of contact clusters as a function of patchy particle model parameters ϵ_{SC} and σ_{SC} at (a) $\epsilon_A = 2.5\ k_BT$, (b) $\epsilon_A = 5.0\ k_BT$, and (c) $\epsilon_A = 7.5\ k_BT$. Mean values of k_{con} are calculated over five independent simulations at each parameter setting and plotted on logarithmic axes to better illuminate the trends. Relative standard errors in k_{con} over the five runs are estimated as 4-29%.

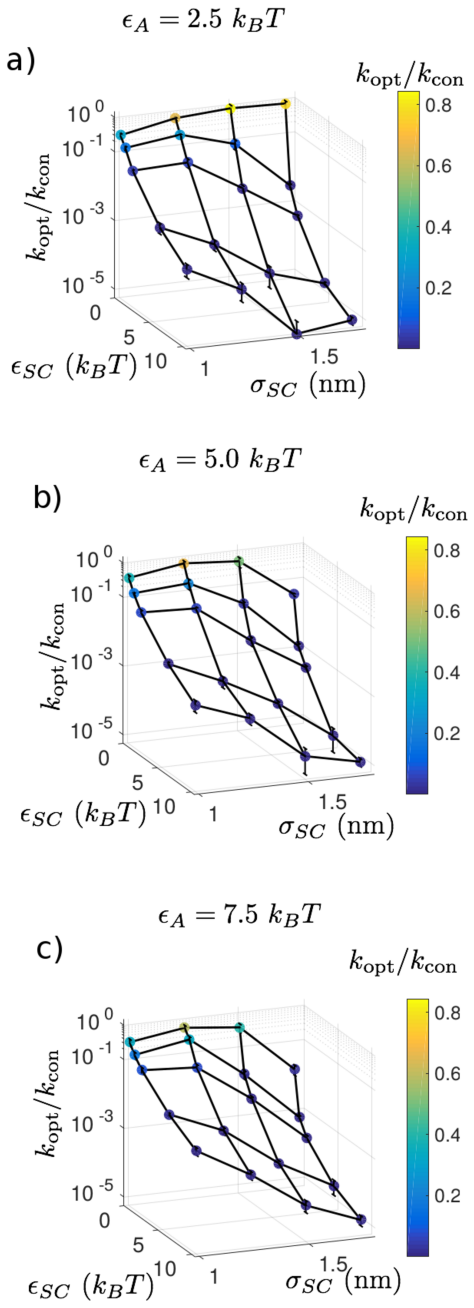


Figure 5: Ratio of the best fit optical and contact cluster growth rates $k_{\mathrm{opt}}/k_{\mathrm{con}}$ as a function of patchy particle model parameters ϵ_{SC} and σ_{SC} at (a) $\epsilon_A = 2.5\ k_BT$, (b) $\epsilon_A = 5.0\ k_BT$, and (c) $\epsilon_A = 7.5\ k_BT$. The ratio is plotted on logarithmic axes to better illuminate the trends.

aggregation driven largely by core–core rather than core–side chain or side chain–side chain interactions, such that further decreases in ϵ_{SC} are significantly more effective in increasing optical cluster rate than those when $\epsilon_{SC} > \epsilon_{BB}$.

The influence of side chain diameter σ_{SC} on the growth rate is less pronounced. In a gross sense, larger values of σ_{SC} tend to elevate k_{opt} for $\epsilon_{SC} < 0.9\ k_BT$ but depress it for $\epsilon_{SC} > 0.9\ k_BT$. However, there is marked non-monotonicity in these trends, and there is evidence for a weak maxima and minima of k_{opt} as a function of σ_{SC} for particular choices of $\{\epsilon_{SC}, \epsilon_A\}$ pairs. This non-monotonic behavior may be understood as the competition of (at least) two effects of increasing side chain diameter: (i) it increases the growth rate by increasing monomer cross-section for productive associative collisions, but (ii) it reduces the range of possible configurations accessible to optical clusters and hampers mutual rearrangements to favor well-aligned core stacking through steric hindrances.

The effect of the cofacial interaction strength ϵ_A is also relatively weak, having relatively minor impact on the shape and vertical shift of the $\epsilon_{SC} - \sigma_{SC}$ surfaces presented in Fig. 3. The only exception to this trend occurs at $\epsilon_{SC} = 0.2\ k_BT$ and $\sigma_{SC} = 1.75$ nm, where dropping ϵ_A from 7.5 k_BT to 2.5 k_BT induces a 5000% increase in k_{opt} from $(2.0 \pm 0.5) \times 10^{-4}$ ns^{-1} to $(1.0 \pm 0.2) \times 10^{-2}$ ns^{-1}. This indicates that in a regime where side chains are weakly interacting and bulky, weaker cofacial interactions are important in promoting good core alignment, likely due to allowing for easier configurational rearrangement.

3.1.3. Small ϵ_{SC} and σ_{SC} promote similar optical and contact cluster growth rates.

Considering Fig. 5, similar trends emerge in the dependence of the optical to contact cluster growth rate upon the model parameters. This parity can be understood because the parameters have a far stronger influence on the growth kinetics for optical clusters than contact clusters, inducing k_{opt} to span a range of five orders of magnitude compared to only one for k_{con} (cf. Figs. 3 and 4). Once again, the side chain interaction strength has the strongest

influence, with small values of ϵ_{SC} promoting commensurate growth rates of optical and contact clusters such that $k_{\text{opt}}/k_{\text{con}}$ approaches unity. Smaller values of the side chain diameter σ_{SC} tend to promote similar optical and contact cluster growth rates, although this trend is muted for $\epsilon_{SC} = 0.2\ k_BT$. The influence of the cofacial interaction strength ϵ_A is relatively weak, except for – as was observed for the absolute optical cluster growth rate – $\epsilon_{SC} = 0.2\ k_BT$ and $\sigma_{SC} = 1.75$ nm, where dropping ϵ_A from 7.5 k_BT to 2.5 k_BT induces a 5000% increase in $k_{\text{opt}}/k_{\text{con}}$ from $(1.6 \pm 0.5) \times 10^{-2}$ to $(8 \pm 2) \times 10^{-1}$.

3.2 Fractal dimension, linearity, and length scales of self-assembled aggregates and networks

In addition to promoting rapid absolute and relative optical cluster growth, it is also desirable to increase the linearity of the self-assembled aggregates. Experimental measurements demonstrate that increased fibril linearity can be correlated with improved optoelectronic properties[53,54], so it is an additional objective for the assembly process that it result in close to 1D supramolecular assemblies over large length scales.

3.2.1. Quantification of fractal dimensionality

We characterize the degree and length scales over which the peptides self-assemble into approximately linear aggregates as a function of the model parameters by calculating the fractal dimension of the system[55,149,150]. We estimate the fractal dimension through a numerical approximation to the correlation integral,

$$C(r) = \lim_{N \to \infty} \frac{g}{N^2}, \tag{6}$$

where g is the number of points separated by a distance less than r. The correlation integral grows with the fractal dimension D of a system as $C(r) \sim r^D$, providing a measure of how the system fills space. We make a discrete approximation to the correlation integral of our

system by using a Heaviside function H[55,150,151],

$$C(\ell) \approx \frac{1}{N(N-1)} \sum_{i,j=1,i\neq j}^{N} H\left(\ell - r_{ij}^{\mathrm{COM}}\right), \tag{7}$$

where r_{ij}^{COM} is the distance between the centers of mass of monomers i and j. By computing $C(\ell)$ at a number of time points over the final 600 μs of each simulation trajectory, we verify that the correlation integral is converged, and that the estimate from the terminal frame is representative of the late-stage self-assembled morphology of the system.

3.2.2. Fractal dimension varies with length scale and model parameters

We present in Fig. 6 the correlation integrals computed from the terminal frame of the five independent simulations conducted at each of 60 different parameter settings. Plotted on log-log axes, the slope of a tangent line to the curve provides an estimate of the fractal dimensionality D as a function of observation length scale ℓ. Collating the data from all calculations makes it challenging to pick out any particular parameter set, but the intention of this plot is to show the relative similarity of the trends in fractal dimension over the wide range of parameters considered. Moreover, the data suggest a natural partitioning of the correlation integral into four distinct regimes. Regimes I and IV have very simple interpretations. Regime I ($\ell \leq 2.5$ nm) is defined by half the linear extent of a single peptide monomer, and the correlation integral, initially zero due to the excluded volume of the peptides, exhibits a sharp jump. Regime IV ($\ell > 121.5$ nm) exhibits a plateau in the correlation integral due to the finite size of the simulation box.

Regime III (30.1 nm $\leq \ell <$ 121.5 nm) corresponds to observations on the scale of tens of peptide lengths, wherein data from all 60 different parameter sets collapse onto a single curve with slope (2.15 $\pm$ 0.06). This observation reveals that the fractal dimension of the self-assembled system at large length scales is insensitive to the particulars of the molecular level chemistry, and the system forms a fractal network with dimensionality $D_{III} = (2.15 \pm$

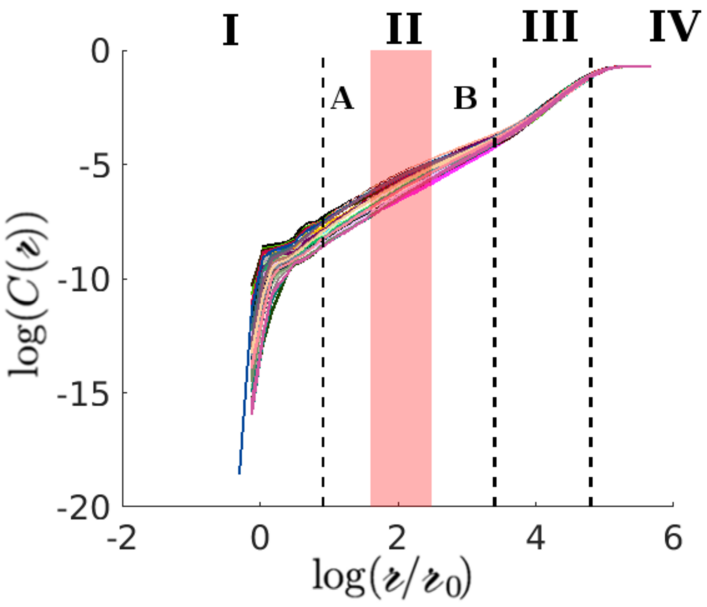


Figure 6: Log of the correlation integral $C(r)$ versus the log of the distance r, which is dedimensionalized by an arbitrary factor $r_0 = 1$ nm. Data are plotted for the terminal frame of the five independent simulations conducted at each of the 60 parameter settings. The data corresponding to each parameter setting are differentiated by color and all plotted together to evince general trends. The dashed black vertical lines demarcate the four different regimes I-IV of the correlation integral described in the text. The shaded red area delimits the range of values over which Regime II is subdivided into Regimes II-A and II-B at each parameter setting.

0.06). Representative snapshots of this network are presented for selected parameter settings in Fig. 7, revealing the hierarchical aggregation of elongated micelles and ribbons into a branched porous network, the nodes of which are formed by the intersection of pseudo-linear supramolecular aggregates. This network is consistent with experimental "matted hair" morphologies of oligopeptide aggregates assembled under quiescent conditions in the absence of aligning hydrodynamic flows that are observed on length scales of $\sim$1 μm[53,70]. The observation of this large-scale network under all parameter settings suggests that its highly conserved formation arises from the intrinsic geometry of the peptide monomers. The structure formed is not space filling due to an entropic and energetic preference for parallel stacking: entropically, elongated rods tend to align to preserve translational entropy at the cost of rotational[97,152,153], and energetically there is a larger number of favorable interactions available along their long axis. Assembly is anticipated to proceed under kinetic control (Section 3.1), so it is expected that these aggregates are not thermodynamic minima but rather morphological states that serve as dynamical attractors for the assembly kinetics of the peptide family under conditions of rapid acidification, and which do not depend sensitively on the chemical details of the oligopeptide monomers. The true thermodynamic minimum is anticipated to be a crystal[133–135], but more sophisticated free energy calculations would be required to resolve this question and quantify the metastability of the observed fractal aggregates[154].

Regime II (2.5 nm $\leq \lambda <$ 30.1 nm) is defined on length scales of about one to ten of peptide diameters and corresponds to the rough length scale of fibril width observed experimentally ($\sim$10 nm)[41,53,70]. The spread in the correlation integral curves within this region indicates that the dimensionality of the self-assembled aggregates on these length scales is influenced by the parameters ϵ_{SC}, σ_{SC}, and ϵ_A. Further analysis shows that all curves within Regime II are better modeled by a two-piece linear fit as opposed to a single linear regression as judged by the Akaike Information Criterion[155]. This analysis reveals that Regime II may actually be divided further into two sub-regions Regime II-A and II-B. We perform the two-piece

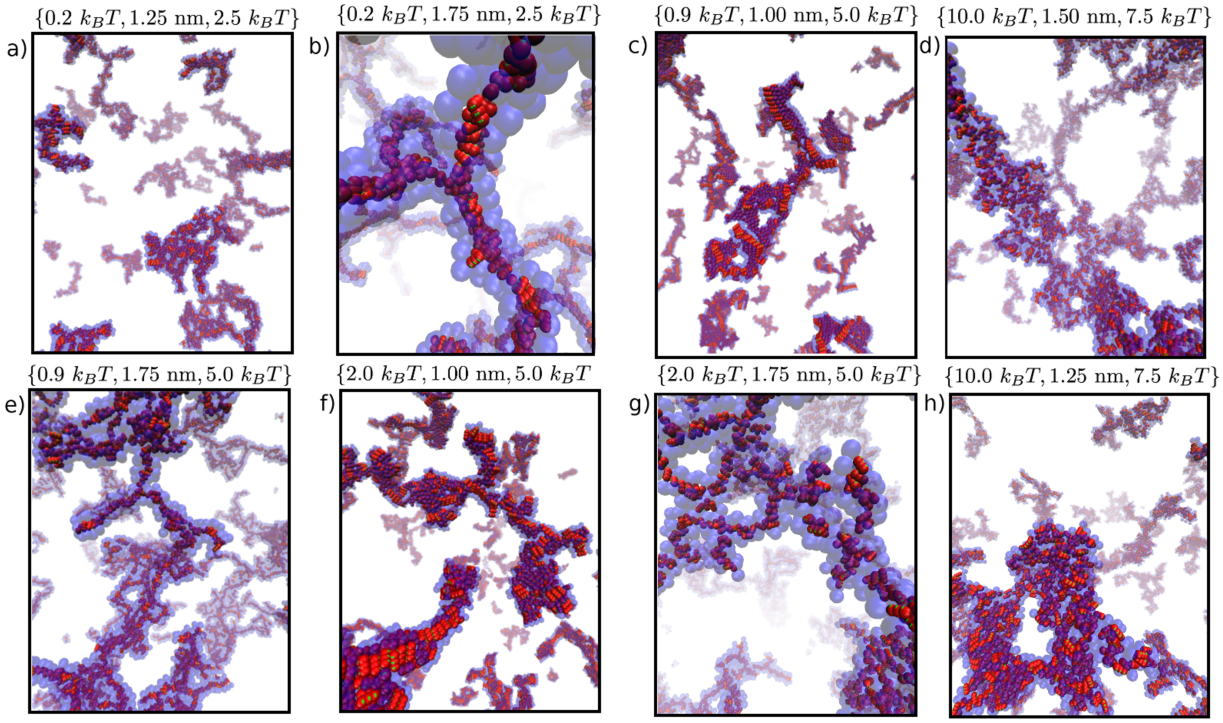


Figure 7: Snapshots of the late-stage self-assembled morphologies for selected $\{\epsilon_{SC}, \sigma_{SC}, \epsilon_A\}$ parameter settings. Side chain (SC) beads are transparent and colored blue, non-cofacial aromatic core (BB) beads are colored red, and cofacial aromatic core (A) beads are colored green. At all parameter settings we observe the formation of a porous fractal network at length scales exceeding ∼30 nm. Each panel is labeled at the top with its $\{\epsilon_{SC}, \sigma_{SC}, \epsilon_A\}$ values. To provide a sense of scale, the red BB beads are shown with diameters of 1 nm corresponding to the minimum of their LJ potential. Snapshots were rendered in VMD[114].

fit and identify the optimal location for the crossover point using the L-method of Salvador and Chan[156]. The precise location of the crossover depends on the parameter set under consideration, but falls within the range 5.0–12.3 nm as indicated by the red shading in Fig. 6.

Regime II-A (2.5 nm $\leq \ell <$ 5.0–12.3 nm) spans length scales between about two and ten core–core stacked peptides. Fig. 8 shows how D_{II-A} varies as a function of the model parameters, spanning a range of 1.4–1.9. As was the case for cluster growth rates, ϵ_A has very little influence upon the fractal dimensionality. This parameter governs the strength of the cofacial interactions that, by construction, mediate parallel stacked linear aggregates, and so this weak dependence is unsurprising. Conversely, the side chain parameters have a relatively large influence upon dimensionality. Large ϵ_{SC} promotes fractal dimensions approaching 2, corresponding to micellar porous networks on a small scale (Fig. 7d,h), whereas small values favor more linear aggregates with dimensionality closer to 1.5 (Fig. 7a-c), corresponding to more core-interacting assemblages. For $\epsilon_{SC} < 2\ k_BT$, both sufficiently small and sufficiently large values of σ_{SC} favor more linear aggregates, but small values of σ_{SC} favor isolated ribbon-like stacks (Fig. 7c), medium values of σ_{SC} favor ribbon-like assemblies that stack side to side (Fig. 7a), and large values of σ_{SC} favor twisted one-dimensional fibrils (Fig. 7b). Notably, there is a moderate anticorrelation between the fractal dimension in Regime II-A and the optical cluster growth rate ($\rho_{\text{Pearson}} = -0.47$, $p < 10^{-6}$).

These observations may be understood as a transition between core-mediated (ribbon-like) and side chain-mediated (porous micellar) interactions. When the side chains are sufficiently sticky (that is when ϵ_{SC} is large enough), core–core interactions are disfavored in comparison to side chain interactions, leading to small core-stacked aggregates of approximately two-three monomers interacting more promiscuously and forming porous micellar structures, similar to the micelles identified by Vácha and Frenkel as one phase of a system of spherocylindrical monomers with attractive end caps[115]. For side chains that are less sticky, a sufficiently small σ_{SC} lessens the likelihood of any side chain interactions, leading to

lower-dimensional ribbon-like structures, similar to the parallel stacking identified by Vácha and Frenkel as one phase of a system of spherocylindrical monomers without attractive end caps[115]. As σ_{SC} approaches σ_{BB}, side chain interactions can lead to more interactions between ribbons, but a sufficiently large σ_{SC} forces the formation of a twisting one-dimensional structure instead of a flat ribbon, which does not interact easily in a flat stack-to-stack manner (compare Fig. 7b with Fig. 7c), leading to a lowered dimensionality. The observed anticorrelation of the growth rate of optical clusters with the dimension of the resulting aggregates demonstrates that the final details of the aggregate structure are at least partially mediated by the kinetics of aggregation, as has also been observed experimentally[54,70].

Regime II-B (5.0–12.3 nm $\leq \lambda <$ 30.1 nm) spans length scales between about ten and thirty core–core stacked monomers. These length scales are larger than the scale of local packing but smaller than the scale of onset of the parameter-independent porous fractal network. The dependence of D_{II-B} on the model parameters is illustrated in Fig. 9, spanning a range of 0.9-1.8. The primary discriminant of the dimensionality of the supramolecular aggregates over these length scales is the side chain diameter σ_{SC}, with small side chains promoting lower-dimensional aggregates (Fig. 7a,c,f,h), and larger side chains favoring assemblies with $D_{II-B} \to 2$ (Fig. 7b,e,g). There is a weak minimum observed for $\epsilon_{SC} = 2\ k_B T$, which becomes mildly less pronounced as ϵ_A increases. Notably, there is a relatively strong correlation between the fractal dimension in Regime II-B and the contact cluster growth rate (Pearson correlation coefficient $\rho_{\text{Pearson}} = 0.82$, $p < 10^{-6}$).

Regime II-B comprises the length scale of transition between the scale of lower-dimensional packing and the scale of the two-dimensional porous network. Systems that grow more slowly are somewhat disconnected on this length scale (cf. Fig. 7a,c) and their dimension is lower because of the gaps in the self-assembled network. Increasing the side chain radius increases the growth rate and D_{II-B} approaches 2 as it increases the size and collisional cross section of the monomers. The mild decrease observed with decreasing ϵ_{SC} corresponds to a mild decrease in the probability of a productive collision for the more side chain-mediated growth

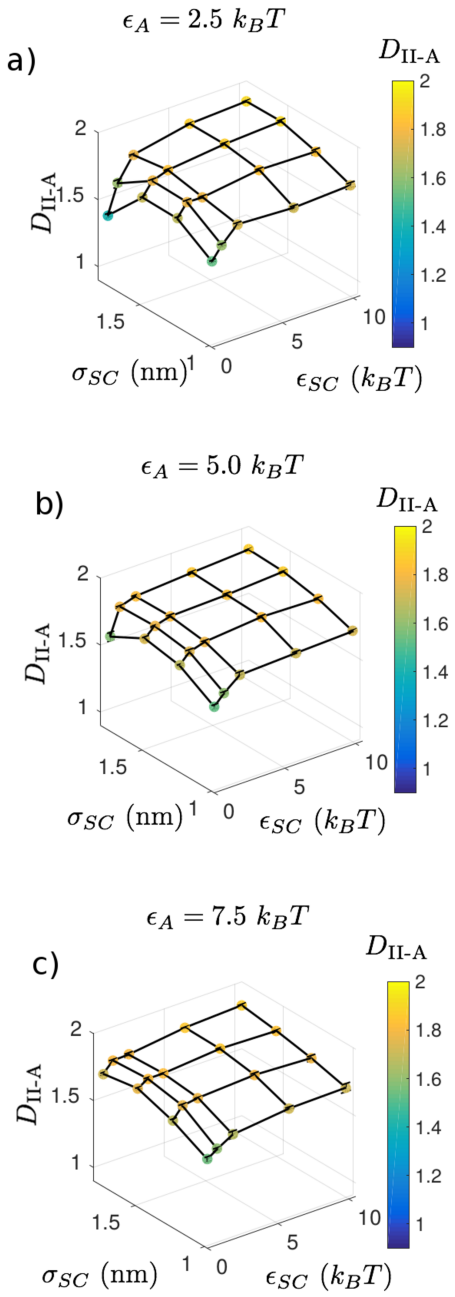


Figure 8: Fractal dimension within Regime II-A as a function of model parameters ϵ_{SC} and σ_{SC} for (a) $\epsilon_A = 2.5\ k_BT$, (b) $\epsilon_A = 5.0\ k_BT$, and (c) $\epsilon_A = 7.5\ k_BT$. Uncertainties in the calculated dimensionality $D_{\text{II-A}}$ are estimated over the terminal frames extracted from five independent simulations.

that occurs when $\epsilon_{SC} > \epsilon_{BB}$. The mild increase observed with decreasing ϵ_{SC} further is attributable to lowering side chain disruptions to primarily core-mediated growth.

3.2.3. The patchy particle model successfully recapitulates properties predicted by the Martini model and measured by experiment

Our model successfully captures aspects of peptide self-assembly previously reported in both experimental and computational studies. Experimentally, the DXXX-OPV3-XXXD and similar systems have been shown to robustly form porous networks of fibrils with widths on the order of tens of nanometers on scales of hundreds of nanometers[24,34,53,70], with the precise morphology controlled by both peptide chemistry[24] and assembly conditions[54,70]. In good agreement with these experimental observations, we observe the robust formation of porous, branched networks with fractal dimension $D \approx 2$ on the order of tens to hundreds of nanometers, and fibril widths on the order of ten nanometers.

The fractal dimension and specific morphologies of these aggregates on shorter length scales is controlled by the specific interactions of the cores and side chains. Prior computational work employing the Martini model that lumps approximately four atoms into each coarse-grained bead showed that the DFAG-OPV3-GAFD system forms amorphous aggregates with dimension ∼1.5 on length scales of tens of nanometers and that there is a hierarchy of optical and contact clusters[55]. The results from the present patchy particle model are in good agreement with these observations, predicting the same hierarchy of optical and contact clusters and the formation of aggregates with fractal dimensionality on these length scales of 0.9–1.9 depending on the precise choice of parameters. The patchy particle model was not parameterized against these higher-resolution simulation results, and the good agreement with the more detailed model provides a validation of its predictive power.

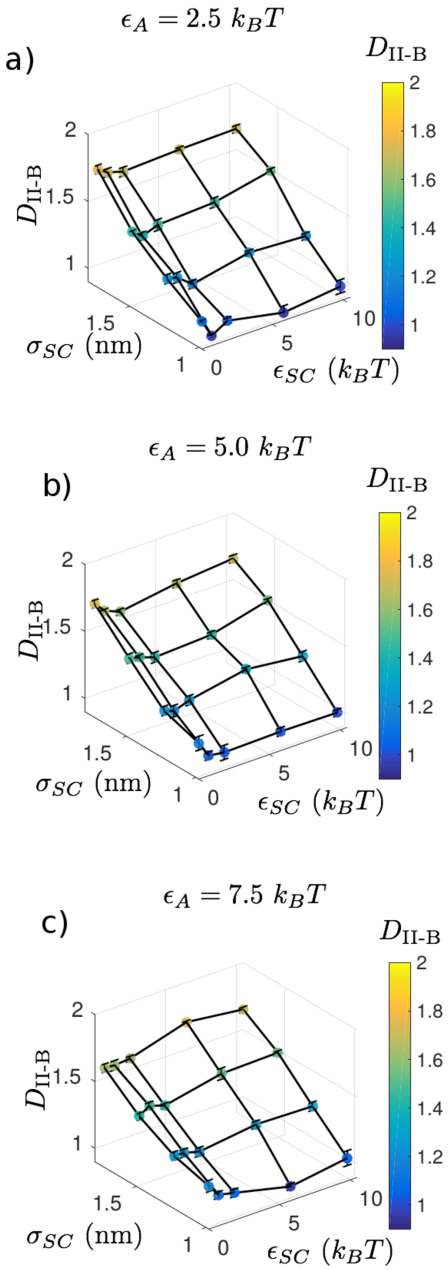


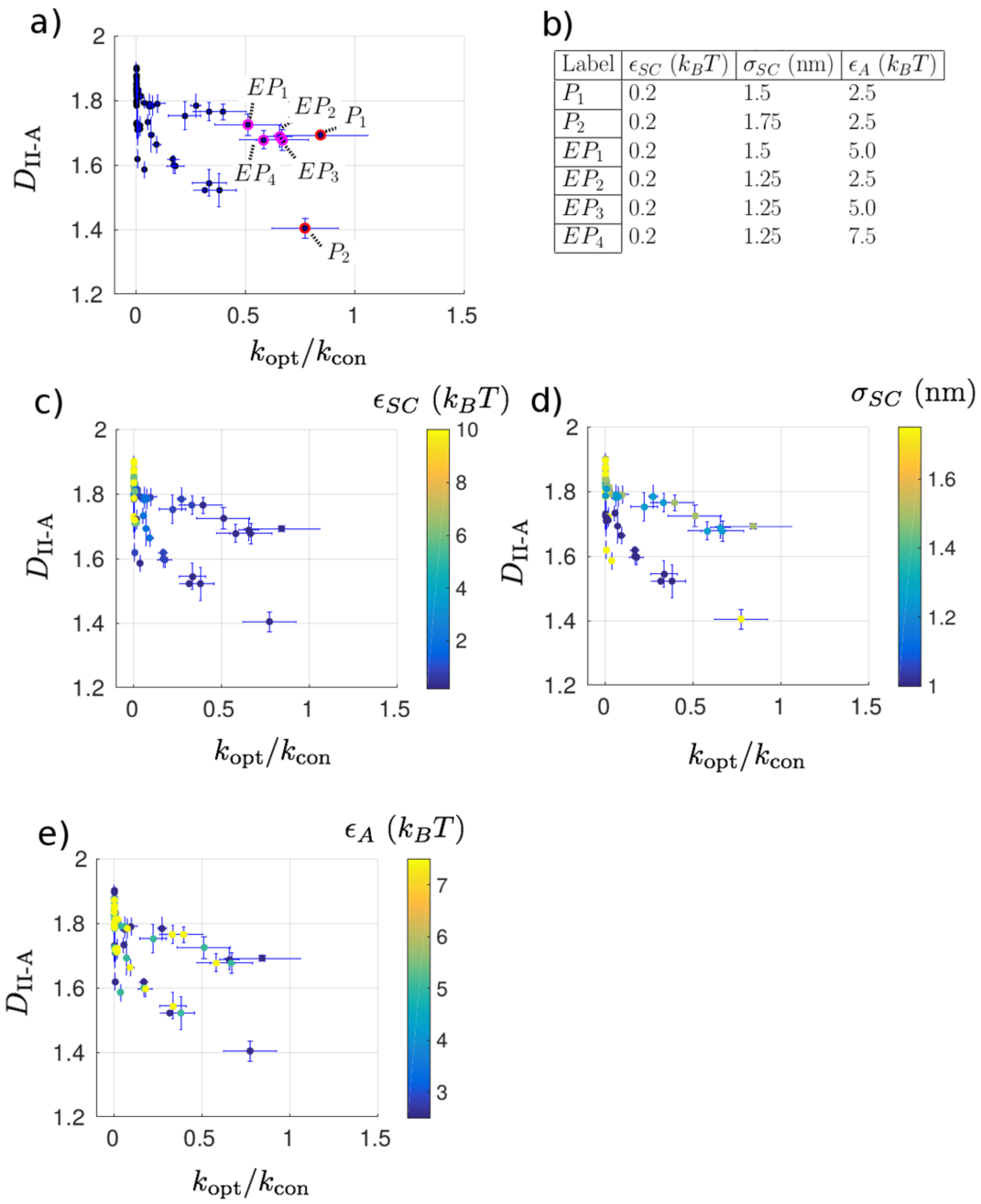
Figure 9: Fractal dimension within Regime II-B as a function of model parameters ϵ_{SC} and σ_{SC} for (a) $\epsilon_A = 2.5\ k_BT$, (b) $\epsilon_A = 5.0\ k_BT$, and (c) $\epsilon_A = 7.5\ k_BT$. Uncertainties in the calculated dimensionality $D_{\text{II-B}}$ are estimated over the terminal frames extracted from five independent simulations.

3.3 Pareto optimization of relative optical cluster growth rate and supramolecular linearity

3.3.1. Identification of Pareto frontier

Having determined the kinetic and morphological properties of assembly under different parameter choices, we now seek to identify the parameter set that (i) maximizes the relative rate of optical to contact cluster growth $k_{\text{opt}}/k_{\text{con}}$ to assure the formation of large optical aggregates, and (ii) minimizes the fractal dimension in Regime II-A D_{II-A} to produce linear fibrils on ∼10 nm length scales that are expected to exhibit good electronic and optical properties due to core–core π stacking. We engage this multi-objective optimization problem by identifying the Pareto frontier within the ensemble of 60 parameter settings explored (Table 1). A point is defined to reside on the Pareto frontier if no other point in the ensemble is superior in all components of the objective function[157]. In the present case, Pareto optimal parameter sets $\{\epsilon_{SC}, \sigma_{SC}, \epsilon_A\}$ are those for which no other parameter set produces both larger $k_{\text{opt}}/k_{\text{con}}$ and smaller D_{II-A}. Pareto frontier points are optimal in the respect that improving any one component of the objective function necessitates a degradation in another. The relative importance of the various components of the objective function may be weighted to ultimately prefer one Pareto optimal point over another.

We present in Fig. 10a a $k_{\text{opt}}/k_{\text{con}}$–$D_{II-A}$ scatter plot for the 60 $\{\epsilon_{SC}, \sigma_{SC}, \epsilon_A\}$ parameter sets. The plot reveals relative growth rate and fibril fractal dimensionality are negatively correlated, possessing a Pearson correlation coefficient of $\rho(k_{opt}/k_{con}, D_{II-A}) = -0.52$ ($p < 10^{-6}$). This connection between kinetics and morphology can be understood as similar optical and contact cluster growth rates producing aggregates with well-aligned aromatic cores that promote linear growth. The Pareto frontier comprises only two points P_1 and P_2, but four additional points EP_1–EP_4 are Pareto proximate in that they have errorbars that overlap with those of either P_1 or P_2. We identify the particular parameter sets corresponding to these six points in Fig. 10b. The Pareto optimal and Pareto proximate points all possess the



Label	ϵ_{SC} (k_BT)	σ_{SC} (nm)	ϵ_A (k_BT)
P_1	0.2	1.5	2.5
P_2	0.2	1.75	2.5
EP_1	0.2	1.5	5.0
EP_2	0.2	1.25	2.5
EP_3	0.2	1.25	5.0
EP_4	0.2	1.25	7.5

Figure 10: Scatter plots of the fractal dimension in Regime II-A, D_{II-A}, versus the relative growth rate of optical clusters with respect to contact clusters, $k_{\text{opt}}/k_{\text{con}}$. Each point reports the calculated pair $(k_{\text{opt}}/k_{\text{con}}, D_{II-A})$ corresponding to each of the 60 $\{\epsilon_{SC}, \sigma_{SC}, \epsilon_A\}$ parameter sets considered. Values are averaged over five independent runs, and errorbars illustrate estimated standard errors. (a) Red circles identify the two Pareto optimal candidates P_1 and P_2, and magenta circles the four candidates EP_1–EP_4 whose error bars overlap those of either P_1 or P_2. (b) Parameter sets $\{\epsilon_{SC}, \sigma_{SC}, \epsilon_A\}$ corresponding to each of the six circled points. Scatter plots colored by (c) ϵ_{SC}, (d) σ_{SC} and (e) ϵ_A to resolve trends in D_{II-A} and $k_{\text{opt}}/k_{\text{con}}$ with respect to these model parameters.

smallest values of $\epsilon_{SC} = 0.2\ k_BT$ studied in this work, but a wide range of σ_{SC} and ϵ_A values. The optimality of small side chain interaction strength ϵ_{SC} may have been anticipated from Figs. 5 and 8 where small ϵ_{SC} led to elevated relative optical to contact growth rates and more linear aggregates within Regime II-A. Coloring the scatter plot by ϵ_{SC} in Fig. 10c clearly illustrates this trend, with increasing ϵ_{SC} away from the Pareto frontier. The side chain size σ_{SC} has a non-monotonic influence on the relative optical cluster growth rate (Fig. 5) and fractal dimension (Fig. 8) at $\epsilon_{SC} = 0.2\ k_BT$. Coloring the scatter plot by σ_{SC} in Fig. 10d fails to resolve any clear trends in this parameter with respect to proximity to the Pareto frontier. As noted previously, ϵ_A has a weak effect on both the relative growth rate and fractal dimensionality, and the scatter plot colored by this model parameter in Fig. 10e fails to show any clear dependency in either component of the objective function.

In sum, the Pareto analysis reveals good regions of parameter space and also provides new understanding of the important determinants of assembly: ϵ_{SC} should be tuned to a small value, the precise value of ϵ_A is unimportant over the range considered, and the dependence on σ_{SC} is relatively complex and non-monotonic. We also note that all Pareto optimal candidates possess $\epsilon_{SC} = 0.2\ k_BT$ at the lowest end of the range considered in this work. Although this range was defined by physically motivated concerns (Section 2.1.1), the results of our analysis suggest that it would likely be worthwhile to extend our investigation to even smaller values of the side chain well depth, and even explore weakly repulsive interactions.

3.3.2. Identification of optimal peptide chemistries

The Pareto analysis identified six patchy particle parameter combinations that produced desirable assembly behaviors. To translate this analysis to peptide design, we now proceed to identify particular oligopeptide chemistries consistent with these parameter sets. Our results have shown ϵ_A to have a rather weak influence on relative growth rate $k_{\text{opt}}/k_{\text{con}}$ (Fig. 5) and fractal dimension D_{II-A} (Fig. 8), so we instead focus on tuning ϵ_{SC} and σ_{SC}. Since the patchy particle model was parameterized for an oligo(p-phenylenevinylene)3 (OPV3)

aromatic core, we can consider the design procedure to be the identification of optimal peptide wings for a DXXX-OPV3-XXXD molecule. We perform the reverse mapping of the patchy particle parameters ϵ_{SC} and σ_{SC} using the techniques detailed in Section 2.1.2: σ_{SC} is estimated from the sums of the SASAs for individual amino acid residues, and ϵ_{SC} for a particular peptide wing sequence is computed from our previously published QSPR model[42].

The QSPR model was developed for peptide wings containing the ten nonpolar amino acid residues at the low pH (pH $\lesssim$ 1) at which assembly proceeds: Ala, Gly, Glu, Ile, Leu, Met, Phe, Trp, Tyr, and Val (Asp residues are excluded as these are reserved to serve as the C-terminal triggers for pH-mediated assembly)[42]. We present in Fig. 11 a ϵ_{SC}-σ_{SC} scatter plot for the $10^3 = 1000$ nonpolar oligopeptide side chain chemistries for which estimates of dimerization free energies are available from the QSPR model (grey crosses). Each point possesses an estimated uncertainty of 3 k_BT in ϵ_{SC} arising from the QSPR model[42] and uncertainties in σ_{SC} propagated from the SASA calculations. We superpose onto this plot the three ϵ_{SC}-σ_{SC} values – (ϵ_{SC},σ_{SC}) = { (0.2 k_BT, 1.25 nm), (0.2 k_BT, 1.5 nm), (0.2 k_BT, 1.75 nm)}– corresponding to the set of two Pareto optimal and four Pareto proximate parameter sets identified from the patchy particle model parameter scan (green crosses). Reverse mapping these optimal parameter sets to particular oligopeptide chemistries is performed by identifying those chemistries for which a Pareto optimal point lies within their error bars. This procedure identifies five particular DXXX-OPV3-XXXD oligopeptide sequences with XXX = {GAG, GGA, GGG, FMI, MFI}. Three oligopeptide chemistries with small amino acids in the peptide wings – GAG, GGA, and GGG – fall near the (ϵ_{SC},σ_{SC}) = (0.2 k_BT, 1.25 nm) Pareto point at the lower edge of physically realizable peptide wing volumes. Two chemistries with somewhat bulkier wings – FMI and MFI – lie near the (ϵ_{SC},σ_{SC}) = (0.2 k_BT, 1.5 nm) point. No chemistries lie near the (ϵ_{SC},σ_{SC}) = (0.2 k_BT, 1.75 nm) point due to the absence of physically realizable peptide sequences with such large volumes.

This analysis has served to map the regions of optimal parameter space identified in the

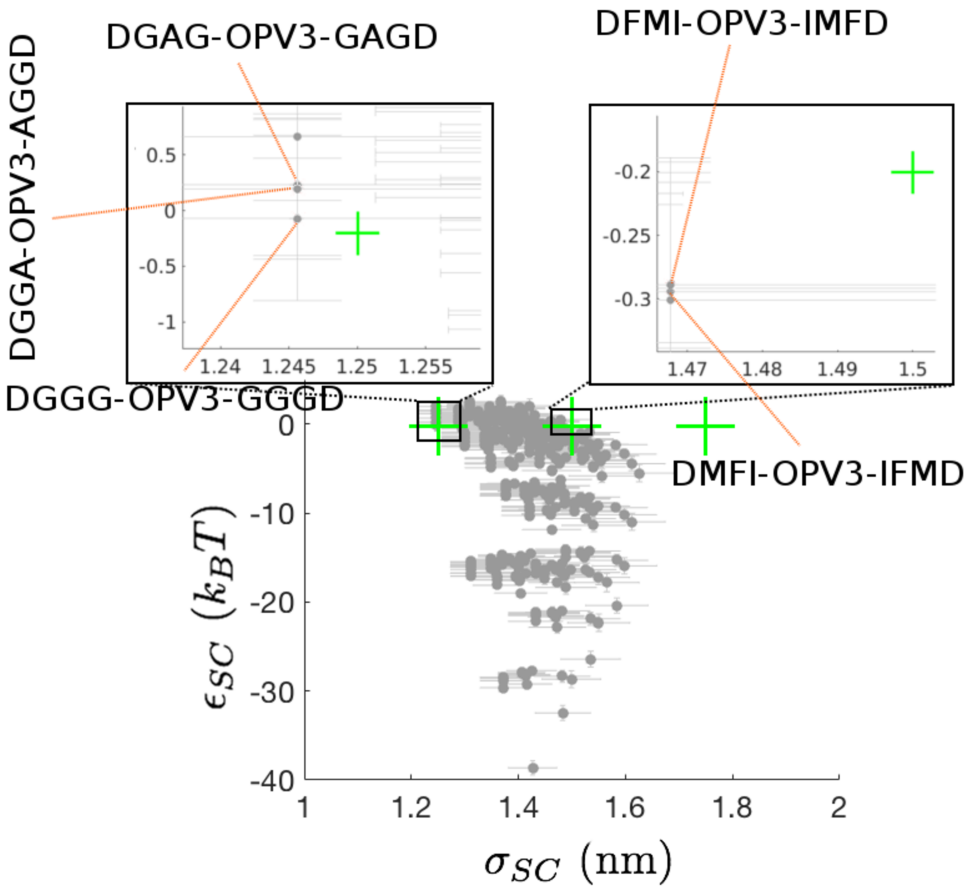


Figure 11: Scatter plot of ϵ_{SC}-σ_{SC} for 1000 nonpolar DXXX-OPV3-XXXD molecules (grey crosses) superposed with the locations of the Pareto optimal and Pareto proximal points defined by the patchy particle parameter sweep (green crosses). Error bars of $\pm 3\ k_BT$ in ϵ_{SC} result from estimated uncertainties in the QSPR model predictions[42], and error bars in σ_{SC} are propagated from uncertainties in the SASA calculations. Five candidates are identified for which a Pareto point lies within their error bars providing a reverse mapping from optimal parameter sets identified in the patchy particle model to particular oligopeptide sequences for future computational or experimental testing.

large-scale patchy particle simulations to particular oligopeptide chemistries that are predicted to possess favorable self-assembly behaviors. The value of the coarse-grained screening and reverse mapping procedure is in presenting an efficient means to direct the search of chemical space towards promising oligopeptide chemistries. In particular, we have identified five candidates for future computational and experimental testing within the 20^3=8000 members of the DXXX-OPV3-XXXD family. Of these candidates, DGAG-OPV3-GAGD has been previously studied and shown to possess desirable optoelectronic properties[24]. The incorporation of Phe residues has been shown to improve aggregation[158], and while many such chemistries have been studied both computationally and experimentally[15,35,41,55,58], neither DFMI nor DMFI has previously been considered. The computational screening and design protocol based on our coarse-grained patchy particle model has therefore identified previously studied oligopeptide chemistries known to possess desirable optoelectronic properties and which were designed based on chemist's intuition, and has also predicted new, and possibly unexpected, promising regions of oligopeptide sequence space.

4. CONCLUSIONS

Inspired by the work of Sărić et al.[80,83] and Zhang et al.[86], we have developed a patchy particle model of the DXXX-Π-XXXD system of optoelectronic self-assembling peptides. Through simplification of the atomistic details and the use of high-performance cyber infrastructure, we were able to directly simulate the self-assembly of ten thousand monomers over hundreds of nanometers and hundreds of microseconds, while still preserving the salient features of the DXXX-Π-XXXD system. The morphological predictions of the patchy model are consistent with experimental measurements and prior molecular simulations employing a higher resolution model, but the model enables access to very long time and length scales with near molecular-level resolution. Furthermore, the inexpensive computational cost of the model enables us to screen over a large parameter space of intermolecular interactions to efficiently identify the important physico-chemical determinants of good assembly behavior and

promising parameter regimes for more detailed computational or experimental investigation.

Our model provides new molecular-level understanding of the key determinants of DXXX-Π-XXXD self-assembly. Decreasing the peptide wing well depth ϵ_{SC} below the well depth of the non-cofacial interactions of aromatic cores ϵ_{BB} leads to significant increase in both absolute and relative contact cluster growth. The strongest determinant of contact cluster growth rate is the excluded volume of the peptide wings, with increasing σ_{SC} corresponding to an increase in the collisional cross section and hence an increase in the overall growth rate. For sufficiently weakly interacting peptide wings, both increasing and decreasing the excluded volume of the wing leads to the formation of linear aggregates with different microscopic morphologies (flat ribbon versus twisted fibril). The small-scale dimensions of aggregates at all parameter sets are correlated with their respective optical cluster growth rates, demonstrating the important connection between kinetics and morphology. On length scales exceeding ∼30 nm, we observe the the formation of an approximately two-dimensional porous network, the dimensionality of which is insensitive to the particular choice of parameters. The assembly rates and resultant morphologies are very weakly dependent on the interaction strength of the parallel stacked interactions ϵ_A over the range of 2.5-7.5 k_BT at $T = 298$ K.

We performed a Pareto optimization in the relative rate of optical to contact cluster growth $k_{\text{opt}}/k_{\text{con}}$ and fractal dimension D_{II-A} of the aggregates over ∼2.5-12 nm length scales, in order to identify $\{\epsilon_{SC}, \sigma_{SC}, \epsilon_A\}$ parameter sets that produced large, linear optical clusters with good core–core alignment that are expected to possess desirable optical and electronic properties. We identified six optimal candidates and interrogated these results to derive insights into the important determinants of aggregation. We found that small ϵ_{SC} values favor good assembly behavior, suggesting that minimizing the interaction strength of the peptide wings – or even exploring weakly repulsive peptide wings – may serve as a precept for rational oligopeptide design. We also uncovered a complex and non-monotonic dependence of assembly quality upon σ_{SC} that is worthy of further study, and an almost

complete insensitivity to ϵ_A. By mapping the Pareto optimal patchy particle models to sequence-defined oligopeptides, we identified five DXXX-OPV3-XXXD chemistries for future computational and experimental study.

Our results also suggest a number of follow-up investigations. First, the Pareto frontier defines a tighter region of $\epsilon_{SC} - \sigma_{SC} - \epsilon_A$ parameter space for more detailed exploration by our patchy particle model. Second, we propose using the five chemistries identified as lying near the Pareto frontier as the starting points for an active learning procedure conducted using higher resolution bead-level models to further refine a pool of candidates for all-atom simulation or experimental testing[159–164]. Third, we propose to perform simulations of aggregation using our patchy particle model under non-equilibrium flow[54,56]. These calculations would allow us to access the length and time scales upon which flow is predicted to have an important influence on aggregation, permit direct comparisons of the observed morphologies with experiment, furnish molecular-level understanding of the coupling between flow and chemistry upon assembly, and perhaps provide new precepts by which to design optimal flow fields.

Overall, this work has explored the interactions of the DXXX-Π-XXXD system at the mesoscale to provide new fundamental understanding of the important molecular determinants of assembly behavior, permit rapid screening over molecular parameter space, and identify good parameter regimes favoring assembly of large, linear optical clusters with well-aligned cores. It provides new rational design principles by which to rationally engineer self-assembling oligopeptides to fabricate supramolecular assemblies for bioelectronic applications, and establishes a new coarse level within a hierarchy of molecular models of varying resolution with which to characterize and engineer these molecules.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. DMR-1729011. This research used resources of the National Energy Research

Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231. This research is part of the Blue Waters sustained-petascale computing project, which is supported by the National Science Foundation (awards OCI-0725070 and ACI-1238993) and the state of Illinois. Blue Waters is a joint effort of the University of Illinois at Urbana-Champaign and its National Center for Supercomputing Applications. RAM gratefully acknowledges a 2017 Blue Waters Graduate Fellowship. We thank Prof. Andela Sărić, Dr. Victor Anisimov, Mike Howard, and Wes Reinhart for fruitful discussions.

References

(1) Rad-Malekshahi, M.; Lempsink, L.; Amidi, M.; Hennink, W. E.; Mastrobattista, E. Biomedical Applications of Self-Assembling Peptides. *Bioconjug. Chem.* **2015**, *27*, 3–18.

(2) French, K. M.; Somasuntharam, I.; Davis, M. E. Self-Assembling Peptide-Based Delivery of Therapeutics for Myocardial Infarction. *Adv. Drug Deliv. Rev.* **2016**, *96*, 40–53.

(3) Dehsorkhi, A.; Castelletto, V.; Hamley, I. W. Self-Assembling Amphiphilic Peptides. *J. Pept. Sci.* **2014**, *20*, 453–467.

(4) Hendricks, M. P.; Sato, K.; Palmer, L. C.; Stupp, S. I. Supramolecular Assembly of Peptide Amphiphiles. *Acc. Chem. Res.* **2017**, *50*, 2440–2448.

(5) Smits, E. C.; Mathijssen, S. G.; Van Hal, P. A.; Setayesh, S.; Geuns, T. C.; Mutsaers, K. A.; Cantatore, E.; Wondergem, H. J.; Werzer, O.; Resel, R. et al. Bottom-Up Organic Integrated Circuits. *Nature* **2008**, *455*, 956.

(6) Pérez, C. M. R.; Stephanopoulos, N.; Sur, S.; Lee, S. S.; Newcomb, C.; Stupp, S. I. The Powerful Functions of Peptide-Based Bioactive Matrices for Regenerative Medicine. *Ann. Biomed. Eng.* **2015**, *43*, 501–514.

(7) Mba, M.; Moretto, A.; Armelao, L.; Crisma, M.; Toniolo, C.; Maggini, M. Synthesis and Self-Assembly of Oligo (p-Phenylenevinylene) Peptide Conjugates in Water. *Chem. Eur. J.* **2011**, *17*, 2044–2047.

(8) Klok, H.-A.; Rösler, A.; Götz, G.; Mena-Osteritz, E.; Bäuerle, P. Synthesis of a Silk-Inspired Peptide–Oligothiophene Conjugate. *Org. Biomol. Chem.* **2004**, *2*, 3541–3544.

(9) Diegelmann, S. R.; Gorham, J. M.; Tovar, J. D. One-Dimensional Optoelectronic

Nanostructures Derived from the Aqueous Self-Assembly of π-Conjugated Oligopeptides. *J. Am. Chem. Soc.* **2008**, *130*, 13840–13841.

(10) Matmour, R.; De Cat, I.; George, S. J.; Adriaens, W.; Leclère, P.; Bomans, P. H.; Sommerdijk, N. A.; Gielen, J. C.; Christianen, P. C.; Heldens, J. T. et al. Oligo (p-Phenylenevinylene)–Peptide Conjugates: Synthesis and Self-Assembly in Solution and at the Solid–Liquid Interface. *J. Am. Chem. Soc.* **2008**, *130*, 14576–14583.

(11) Stone, D. A.; Hsu, L.; Stupp, S. I. Self-Assembling Quinquethiophene–Oligopeptide Hydrogelators. *Soft Matter* **2009**, *5*, 1990–1993.

(12) Burroughes, J.; Bradley, D.; Brown, A.; Marks, R.; Mackay, K.; Friend, R.; Burns, P.; Holmes, A. Light-Emitting Diodes Based on Conjugated Polymers. *Nature* **1990**, *347*, 539.

(13) Tovar, J. D. Supramolecular Construction of Optoelectronic Biomaterials. *Acc. Chem. Res.* **2013**, *46*, 1527–1537.

(14) Tovar, J. D. Photon Management in Supramolecular Peptide Nanomaterials. *Bioinspiration Biomim.* **2017**, *13*, 015004.

(15) Liyanage, W.; Ardoña, H. A. M.; Mao, H.-Q.; Tovar, J. D. Cross-Linking Approaches to Tuning the Mechanical Properties of Peptide π-Electron Hydrogels. *Bioconjug. Chem.* **2016**, *28*, 751–759.

(16) Panda, S. S.; Katz, H. E.; Tovar, J. D. Solid-State Electrical Applications of Protein and Peptide Based Nanomaterials. *Chem. Soc. Rev.* **2018**, *47*, 3640–3658.

(17) Pinotsi, D.; Grisanti, L.; Mahou, P.; Gebauer, R.; Kaminski, C. F.; Hassanali, A.; Schierle, G. S. K. Proton Transfer and Structure-Specific Fluorescence in Hydrogen Bond-Rich Protein Structures. *J. Am. Chem. Soc.* **2016**, *138*, 3046–3057.

(18) Handelman, A.; Kuritz, N.; Natan, A.; Rosenman, G. Reconstructive Phase Transition in Ultrashort Peptide Nanostructures and Induced Visible Photoluminescence. *Langmuir* **2016**, *32*, 2847–2862.

(19) Guo, X.; Baumgarten, M.; Müllen, K. Designing π-Conjugated Polymers for Organic Electronics. *Prog. Polym. Sci.* **2013**, *38*, 1832–1908.

(20) Bian, L.; Zhu, E.; Tang, J.; Tang, W.; Zhang, F. Recent Progress in the Design of Narrow Bandgap Conjugated Polymers for High-Efficiency Organic Solar Cells. *Prog. Polym. Sci.* **2012**, *37*, 1292–1331.

(21) Kim, S. H.; Parquette, J. R. A Model for the Controlled Assembly of Semiconductor Peptides. *Nanoscale* **2012**, *4*, 6940–6947.

(22) Hoeben, F. J. M.; Jonkheijm, P.; Meijer, E. W.; Schenning, A. P. H. J. About Supramolecular Assemblies of π-Conjugated Systems. *Chem. Rev.* **2005**, *105*, 1491–1546.

(23) Vadehra, G. S.; Wall, B. D.; Diegelmann, S. R.; Tovar, J. D. On-Resin Dimerization Incorporates a Diverse Array of π-Conjugated Functionality within Aqueous Self-Assembling Peptide Backbones. *Chem. Commun.* **2010**, *46*, 3947–3949.

(24) Wall, B. D.; Zacca, A. E.; Sanders, A. M.; Wilson, W. L.; Ferguson, A. L.; Tovar, J. D. Supramolecular Polymorphism: Tunable Electronic Interactions within π-Conjugated Peptide Nanostructures Dictated by Primary Amino Acid Sequence. *Langmuir* **2014**, *30*, 5946–5956.

(25) Mitschke, U.; Bäuerle, P. The Electroluminescence of Organic Materials. *J. Mater. Chem.* **2000**, *10*, 1471–1507.

(26) Roncali, J. Conjugated Poly(thiophenes): Synthesis, Functionalization, and Applications. *Chem. Rev.* **1992**, *92*, 711–738.

(27) Fichou, D. *Handbook of Oligo-and Polythiophenes*; Weinheim: Wiley VCH, 1999.

(28) Newman, C. R.; Frisbie, C. D.; da Silva Filho, D. A.; Brédas, J.-L.; Ewbank, P. C.; Mann, K. R. Introduction to Organic Thin Film Transistors and Design of n-Channel Organic Semiconductors. *Chem. Mater.* **2004**, *16*, 4436–4451.

(29) Hoppe, H.; Sariciftci, N. S. In *Photoresponsive Polymers II*; Marder, S., Lee, K.-S., Eds.; Springer, 2008; pp 1–86.

(30) Beaujuge, P. M.; Reynolds, J. R. Color Control in π-Conjugated Organic Polymers for Use in Electrochromic Devices. *Chem. Rev.* **2010**, *110*, 268–320.

(31) Marty, R.; Szilluweit, R.; Sánchez-Ferrer, A.; Bolisetty, S.; Adamcik, J.; Mezzenga, R.; Spitzner, E.-C.; Feifer, M.; Steinmann, S. N.; Corminboeuf, C. et al. Hierarchically Structured Microfibers of "Single Stack" Perylene Bisimide and Quaterthiophene Nanowires. *ACS Nano* **2013**, *7*, 8498–8508.

(32) Ardoña, H. A. M.; Besar, K.; Togninalli, M.; Katz, H. E.; Tovar, J. D. Sequence-Dependent Mechanical, Photophysical and Electrical Properties of pi-Conjugated Peptide Hydrogelators. *J. Mater. Chem. C* **2015**, *3*, 6505–6514.

(33) Sanders, A. M.; Magnanelli, T. J.; Bragg, A. E.; Tovar, J. D. Photoinduced Electron Transfer within Supramolecular Donor–Acceptor Peptide Nanostructures under Aqueous Conditions. *J. Am. Chem. Soc.* **2016**, *138*, 3362–3370.

(34) Ardoña, H. A. M.; Tovar, J. D. Energy Transfer within Responsive π-Conjugated Coassembled Peptide-based Nanostructures in Aqueous Environments. *Chem. Sci.* **2015**, *6*, 1474–1484.

(35) Thurston, B. A.; Tovar, J. D.; Ferguson, A. L. Thermodynamics, Morphology, and Kinetics of Early-stage Self-Assembly of π-Conjugated Oligopeptides. *Mol. Sim.* **2016**, 1–21.

(36) Ardoña, H. A. M.; Tovar, J. D. Peptide π-Electron Conjugates: Organic Electronics for Biology? *Bioconjug. Chem.* **2015**, *26*, 2290–2302.

(37) Wall, B. D.; Diegelmann, S. R.; Zhang, S.; Dawidczyk, T. J.; Wilson, W. L.; Katz, H. E.; Mao, H.-Q.; Tovar, J. D. Aligned Macroscopic Domains of Optoelectronic Nanostructures Prepared via Shear-Flow Assembly of Peptide Hydrogels. *Adv. Mater.* **2011**, *23*, 5009–5014.

(38) Sanders, A. M.; Tovar, J. D. Solid-phase Pd-Catalysed Cross-coupling Methods for the Construction of π-Conjugated Peptide Nanomaterials. *Supramol. Chem.* **2014**, *26*, 259–266.

(39) Sanders, A. M.; Kale, T. S.; Katz, H. E.; Tovar, J. D. Solid-Phase Synthesis of Self-Assembling Multivalent π-Conjugated Peptides. *ACS Omega* **2017**, *2*, 409–419.

(40) Rasale, D. B.; Maity, I.; Das, A. K. Lipase Catalyzed Inclusion of Gastrodigenin for the Evolution of Blue Light Emitting Peptide Nanofibers. *Chem. Commun.* **2014**, *50*, 8685–8688.

(41) Wall, B. D.; Zhou, Y.; Mei, S.; Ardoña, H. A. M.; Ferguson, A. L.; Tovar, J. D. Variation of Formal Hydrogen-Bonding Networks within Electronically Delocalized π-Conjugated Oligopeptide Nanostructures. *Langmuir* **2014**, *30*, 11375–11385.

(42) Thurston, B. A.; Ferguson, A. L. Machine Learning and Molecular Design of Self-Assembling π-Conjugated Oligopeptides. *Mol. Sim.* **2018**, *44*, 1–16.

(43) Fleming, S.; Ulijn, R. V. Design of Nanostructures Based on Aromatic Peptide Amphiphiles. *Chem. Soc. Rev.* **2014**, *43*, 8150–8177.

(44) Zelzer, M.; Ulijn, R. V. Next-Generation Peptide Nanomaterials: Molecular Networks, Interfaces and Supramolecular Functionality. *Chem. Soc. Rev.* **2010**, *39*, 3351–3357.

(45) Babu, S. S.; Praveen, V. K.; Ajayaghosh, A. Functional π-Gelators and Their Applications. *Chem. Rev.* **2014**, *114*, 1973–2129.

(46) Kale, T. S.; Marine, J. E.; Tovar, J. D. Self-Assembly and Associated Photophysics of Dendron-Appended Peptide-π-Peptide Triblock Macromolecules. *Macromolecules* **2017**, *50*, 5315–5322.

(47) Konda, M.; Maiti, S.; Jadhav, R. G.; Das, A. K. Redox-Active Peptide-Functionalized Quinquethiophene-Based Electrochromic π-Gel. *Chem.–Asian J.* **2018**, *13*, 204–209.

(48) Ardoña, H. A. M.; Draper, E. R.; Citossi, F.; Wallace, M.; Serpell, L. C.; Adams, D. J.; Tovar, J. D. Kinetically Controlled Coassembly of Multichromophoric Peptide Hydrogelators and the Impacts on Energy Transport. *J. Am. Chem. Soc.* **2017**, *139*, 8685–8692.

(49) Digennaro, A.; Wennemers, H.; Joshi, G.; Schmid, S.; Mena-Osteritz, E.; Bäuerle, P. Chiral Suprastructures of Asymmetric Oligothiophene-Hybrids Induced by a Single Proline. *Chem. Commun.* **2013**, *49*, 10929–10931.

(50) Yardeni, J. L.; Amit, M.; Ashkenasy, G.; Ashkenasy, N. Sequence Dependent Proton Conduction in Self-Assembled Peptide Nanostructures. *Nanoscale* **2016**, *8*, 2358–2366.

(51) Schillinger, E.-K.; Kümin, M.; Digennaro, A.; Mena-Osteritz, E.; Schmid, S.; Wennemers, H.; Bäuerle, P. Guiding Suprastructure Chirality of an Oligothiophene by a Single Amino Acid. *Chem. Mater.* **2013**, *25*, 4511–4521.

(52) López-Andarias, J.; Rodriguez, M. J.; Atienza, C.; López, J. L.; Mikie, T.; Casado, S.; Seki, S.; Carrascosa, J. L.; Martín, N. Highly Ordered n/p-Co-Assembled Materials with Remarkable Charge Mobilities. *J. Am. Chem. Soc.* **2015**, *137*, 893–897.

(53) Wall, B. D.; Tovar, J. D. Synthesis and Characterization of π-Conjugated Peptide-Based Supramolecular Materials. *Pure Appl. Chem.* **2012**, *84*, 1039–1045.

(54) Marciel, A. B.; Tanyeri, M.; Wall, B. D.; Tovar, J. D.; Schroeder, C. M.; Wilson, W. L. Fluidic-Directed Assembly of Aligned Oligopeptides with π-Conjugated Cores. *Adv. Mater.* **2013**, *25*, 6398–6404.

(55) Mansbach, R. A.; Ferguson, A. L. Coarse-Grained Molecular Simulation of the Hierarchical Self-Assembly of π-Conjugated Optoelectronic Peptides. *J. Phys. Chem. B* **2017**, 1684–1706.

(56) Mansbach, R.; Ferguson, A. Control of the Hierarchical Assembly of π-Conjugated Optoelectronic Peptides by pH and Flow. *Org. Biomol. Chem.* **2017**, *15*, 5484–5502.

(57) Lampel, A.; Ulijn, R. V.; Tuttle, T. Guiding Principles for Peptide Nanotechnology through Directed Discovery. *Chem. Soc. Rev.* **2018**, 3737–3758.

(58) Valverde, L. R.; Thurston, B. A.; Ferguson, A. L.; Wilson, W. L. Evidence for Prenucleated Fibrilogenesis of Acid-Mediated Self-Assembling Oligopeptides via Molecular Simulation and Fluorescence Correlation Spectroscopy. *Langmuir* **2018**, *35*, 7346–7354.

(59) Ulman, A. *An Introduction to Ultrathin Organic Films: From Langmuir–Blodgett to Self–Assembly*; Academic Press, 2013.

(60) Shen, X.-X.; Han, G.-C.; Yi, Y.-P. Multiscale Description of Molecular Packing and Electronic Processes in Small-Molecule Organic Solar Cells. *Chin. Chem. Lett.* **2016**, *27*, 1453–1463.

(61) Cohen, E.; Weissman, H.; Pinkas, I.; Shimoni, E.; Rehak, P.; Kral, P.; Rybtchinski, B. Controlled Self-Assembly of Photofunctional Supramolecular Nanotubes. *ACS Nano* **2017**, *12*, 317–326.

(62) Liu, X.; Wei, Z.; Balla, I.; Mannix, A. J.; Guisinger, N. P.; Luijten, E.; Hersam, M. C.

Self-Assembly of Electronically Abrupt Borophene/Organic Lateral Heterostructures. *Sci. Adv.* **2017**, *3*, e1602356.

(63) Zhao, Y.; Wu, Q.; Chen, Q.; Wang, J. Molecular Self-Assembly on Two-Dimensional Atomic Crystals: Insights from Molecular Dynamics Simulations. *J. Phys. Chem. Lett.* **2015**, *6*, 4518–4524.

(64) Palczynski, K.; Heimel, G.; Heyda, J.; Dzubiella, J. Growth and Characterization of Molecular Crystals of Para-Sexiphenyl by All-Atom Computer Simulations. *Cryst. Growth Des.* **2014**, *14*, 3791–3799.

(65) Shinoda, K.; Shinoda, W.; Liew, C. C.; Tsuzuki, S.; Morikawa, Y.; Mikami, M. Two-Dimensional Self-Assembled Structures of Adenine Molecules: Modeling and Simulation. *Surf. Sci.* **2004**, *556*, 109–120.

(66) Zhou, Y.; Li, B.; Li, S.; Ardoña, H. A. M.; Wilson, W. L.; Tovar, J. D.; Schroeder, C. M. Concentration-Driven Assembly and Sol–Gel Transition of π-Conjugated Oligopeptides. *ACS Cent. Sci.* **2017**, *3*, 986–994.

(67) Li, B.; Valverde, L. R.; Zhang, F.; Zhou, Y.; Li, S.; Diao, Y.; Wilson, W. L.; Schroeder, C. M. Macroscopic Alignment and Assembly of π-Conjugated Oligopeptides Using Colloidal Microchannels. *ACS Appl. Mater. Interfaces* **2017**, *9*, 41586–41593.

(68) Besar, K.; Ardoña, H. A. M.; Tovar, J. D.; Katz, H. E. Demonstration of Hole Transport and Voltage Equilibration in Self-Assembled π-Conjugated Peptide Nanostructures Using Field-Effect Transistor Architectures. *ACS Nano* **2015**, *9*, 12401–12409.

(69) Sanders, A. M.; Dawidczyk, T. J.; Katz, H. E.; Tovar, J. D. Peptide-Based Supramolecular Semiconductor Nanomaterials via Pd-Catalyzed Solid-Phase "Dimerizations". *ACS Macro Lett.* **2012**, *1*, 1326–1329.

(70) Li, B.; Li, S.; Zhou, Y.; Ardoña, H. A. M.; Valverde, L. R.; Wilson, W. L.; Tovar, J. D.; Schroeder, C. M. Nonequilibrium Self-Assembly of π-Conjugated Oligopeptides in Solution. *ACS Appl. Mater. Interfaces* **2017**, *9*, 3977–3984.

(71) Noro, M. G.; Frenkel, D. Extended Corresponding-States Behavior for Particles with Variable Range Attractions. *J. Chem. Phys.* **2000**, *113*, 2941–2944.

(72) Foffi, G.; Sciortino, F. On the Possibility of Extending the Noro–Frenkel Generalized Law of Correspondent States to Nonisotropic Patchy Interactions. *J. Phys. Chem. B* **2007**, *111*, 9702–9705.

(73) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. Essential Dynamics of Proteins. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 412–425.

(74) Vakser, I. A. Low-resolution Structural Modeling of Protein Interactome. *Curr. Opin. Struct. Biol.* **2013**, *23*, 198–205.

(75) Wu, C.; Shea, J.-E. Coarse-Grained Models for Protein Aggregation. *Curr. Opin. Struct. Biol.* **2011**, *21*, 209–220.

(76) Head-Gordon, T.; Brown, S. Minimalist Models for Protein Folding and Design. *Curr. Opin. Struct. Biol.* **2003**, *13*, 160–167.

(77) Bieler, N. S.; Knowles, T. P. J.; Frenkel, D.; Vácha, R.; Devlin, G. Connecting Macroscopic Observables and Microscopic Assembly Events in Amyloid Formation Using Coarse Grained Simulations. *PLoS Comput. Biol.* **2012**, *8*, e1002692.

(78) Ranganathan, S.; Maji, S. K.; Padinhateeri, R. Defining a Physical Basis for Diversity in Protein Self-Assemblies Using a Minimal Model. *J. Am. Chem. Soc.* **2016**, *138*, 13911–13922.

(79) Di Michele, L.; Eiser, E.; Foderà, V. Minimal Model for Self-Catalysis in the Formation of Amyloid-Like Elongated Fibrils. *J. Phys. Chem. Lett.* **2013**, *4*, 3158–3164.

(80) Šarić, A.; Buell, A. K.; Meisl, G.; Michaels, T.; Dobson, C. M.; Linse, S.; Knowles, T.; Frankel, D. Physical Determinants of the Self-Replication of Protein Fibrils. *Nat. Phys.* **2016**, *12*, 874–880.

(81) Kern, N.; Frenkel, D. Fluid–Fluid Coexistence in Colloidal Systems with Short-Ranged Strongly Directional Attraction. *J. Chem. Phys.* **2003**, *118*, 9882–9889.

(82) Vácha, R.; Linse, S.; Lund, M. Surface Effects on Aggregation Kinetics of Amyloidogenic Peptides. *J. Am. Chem. Soc.* **2014**, *136*, 11776–11782.

(83) Šarić, A.; Chebaro, Y. C.; Knowles, T. P. J.; Frenkel, D. Crucial Role of Nonspecific Interactions in Amyloid Nucleation. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 17869–17874.

(84) Vácha, R.; Frenkel, D. Simulations Suggest Possible Novel Membrane Pore Structure. *Langmuir* **2014**, *30*, 1304–1310.

(85) Kabelka, I.; Vácha, R. Optimal Conditions for Opening of Membrane Pore by Amphiphilic Peptides. *J. Chem. Phys.* **2015**, *143*, 243115.

(86) Zhang, Z.; Glotzer, S. C. Self-Assembly of Patchy Particles. *Nano Lett.* **2004**, *4*, 1407–1413.

(87) Teixeira, P.; Tavares, J. Phase Behaviour of Pure and Mixed Patchy Colloids – Theory and Simulation. *Curr. Opin. Colloid Interface Sci.* **2017**, *30*, 16–24.

(88) Bianchi, E.; Largo, J.; Tartaglia, P.; Zaccarelli, E.; Sciortino, F. Phase Diagram of Patchy Colloids: Towards Empty Liquids. *Phys. Rev. Lett.* **2006**, *97*, 168301.

(89) Zaccarelli, E. Colloidal Gels: Equilibrium and Non-Equilibrium Routes. *J. Phys. Condens. Matter* **2007**, *19*, 323101.

(90) Sciortino, F.; Giacometti, A.; Pastore, G. Phase Diagram of Janus Particles. *Phys. Rev. Lett.* **2009**, *103*, 237801.

(91) Doye, J. P. K.; Louis, A. A.; Lin, I.-C.; Allen, L. R.; Noya, E. G.; Wilber, A. W.; Kok, H. C.; Lyus, R. Controlling Crystallization and its Absence: Proteins, Colloids and Patchy Models. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2197.

(92) Dorsaz, N.; Filion, L.; Smallenburg, F.; Frenkel, D. Spiers Memorial Lecture: Effect of Interaction Specificity on the Phase Behaviour of Patchy Particles. *Faraday Discuss.* **2012**, *159*, 9.

(93) Gögelein, C.; Nägele, G.; Tuinier, R.; Gibaud, T.; Stradner, A.; Schurtenberger, P. A Simple Patchy Colloid Model for the Phase Behavior of Lysozyme Dispersions. *J. Chem. Phys.* **2008**, *129*, 085102.

(94) Liu, H.; Kumar, S. K.; Sciortino, F. Vapor-Liquid Coexistence of Patchy Models: Relevance to Protein Phase Behavior. *J. Chem. Phys.* **2007**, *127*, 084902.

(95) Mao, R.; Lee, M.-T.; Vishnyakov, A.; Neimark, A. V. Modeling Aggregation of Ionic Surfactants Using a Smeared Charge Approximation in Dissipative Particle Dynamics Simulations. *J. Phys. Chem. B* **2015**, *119*, 11673–11683.

(96) Damasceno, P. F.; Engel, M.; Glotzer, S. C. Crystalline Assemblies and Densest Packings of a Family of Truncated Tetrahedra and the Role of Directional Entropic Forces. *ACS Nano* **2012**, *6*, 609–614.

(97) van Anders, G.; Ahmed, N. K.; Smith, R.; Engel, M.; Glotzer, S. C. Entropically Patchy Particles: Engineering Valence through Shape Entropy. *ACS Nano* **2014**, *8*, 931–940.

(98) Newton, A. C.; Groenewold, J.; Kegel, W. K.; Bolhuis, P. G. Rotational Diffusion Affects the Dynamical Self-Assembly Pathways of Patchy Particles. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 15308–15313.

(99) Giacometti, A.; Lado, F.; Largo, J.; Pastore, G.; Sciortino, F. Effects of Patch Size and Number within a Simple Model of Patchy Colloids. *J. Chem. Phys.* **2010**, *132*, 174110.

(100) Bianchi, E.; Tartaglia, P.; Zaccarelli, E.; Sciortino, F. Theoretical and Numerical Study of the Phase Diagram of Patchy Colloids: Ordered and Disordered Patch Arrangements. *J. Chem. Phys.* **2008**, *128*, 144504.

(101) Romano, F.; Sanz, E.; Sciortino, F. Phase Diagram of a Tetrahedral Patchy Particle Model for Different Interaction Ranges. *J. Chem. Phys.* **2010**, *132*, 184501.

(102) Romano, F.; Sciortino, F. Two Dimensional Assembly of Triblock Janus Particles into Crystal Phases in the Two Bond per Patch Limit. *Soft Matter* **2011**, *7*, 5799.

(103) Noya, E.; Vega, C. Phase Diagram of Model Anisotropic Particles with Octahedral Symmetry. *J. Chem. Phys.* **2007**, *127*, 054501.

(104) Lomakin, A.; Asherie, N.; Benedek, G. B. Aeolotopic Interactions of Globular Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9465–9468.

(105) Hloucha, M.; Lodge, J.; Lenhoff, A.; Sandler, S. A Patch–Antipatch Representation of Specific Protein Interactions. *J. Cryst. Growth* **2001**, *232*, 195–203.

(106) Morgan, J. W. R.; Chakrabarti, D.; Dorsaz, N.; Wales, D. J. Designing a Bernal Spiral from Patchy Colloids. *ACS Nano* **2013**, *7*, 1246–1256.

(107) Long, A. W.; Ferguson, A. L. Nonlinear Machine Learning of Patchy Colloid Self-Assembly Pathways and Mechanisms. *J. Phys. Chem. B* **2014**, *118*, 4228–4244.

(108) Long, A. W.; Zhang, J.; Granick, S.; Ferguson, A. L.; Debenedetti, P. G.; Link, A. J.; Zucker, S. W.; Yu, Z.; Kalanyan, B.; Parsons, G. N. et al. Machine Learning Assembly Landscapes from Particle Tracking Data. *Soft Matter* **2015**, *11*, 8141–8153.

(109) Long, A. W.; Phillips, C. L.; Jankowksi, E.; Ferguson, A. L. Nonlinear Machine Learning and Design of Reconfigurable Digital Colloids. *Soft Matter* **2016**, *12*, 7119–7135.

(110) Long, A. W.; Ferguson, A. L. Rational Design of Patchy Colloids via Landscape Engineering. *Mol. Syst. Des. Eng.* **2018**, *3*, 49–65.

(111) Xu, J.; Wang, Y.; He, X.; Gunton, J. D.; Brett, G.; Yu, M.; Xia, L.; Ji, Q.; Meng, L.; Song, W. et al. Self-Assembly of Janus Ellipsoids: a Brownian Dynamics Simulation with a Quantitative Nonspherical-Particle Model. *Soft Matter* **2015**, *11*, 7433–7439.

(112) Liu, Y.; Li, W.; Perez, T.; Gunton, J. D.; Brett, G. Self Assembly of Janus Ellipsoids. *Langmuir* **2012**, *28*, 3–9.

(113) Carpency, T. N.; Gunton, J. D.; Rickman, J. M. Phase Behavior of Patchy Spheroidal Fluids. *J. Chem. Phys.* **2016**, *145*, 214904.

(114) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.

(115) Vácha, R.; Frenkel, D. Relation Between Molecular Shape and the Morphology of Self-Assembling Aggregates: A Simulation Study. *Biophys. J.* **2011**, *101*, 1432–1439.

(116) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed.; Academic Press: San Diego, 2001.

(117) van Anders, G.; Klotsa, D.; Ahmed, N. K.; Engel, M.; Glotzer, S. C. Understanding Shape Entropy through Local Dense Packing. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 4812–21.

(118) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J. The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.* **2008**, *4*, 819–834.

(119) de Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schäfer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved Parameters for the Martini Coarse-Grained Protein Force Field. *J. Chem. Theory Comput.* **2013**, *9*, 687–97.

(120) Marrink, S. J.; Tieleman, D. P. Perspective on the Martini Model. *Chem. Soc. Rev.* **2013**, *42*.

(121) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

(122) Yesylevskyy, S. O.; Schäfer, L. V.; Sengupta, D.; Marrink, S. J. Polarizable Water Model for the Coarse-Grained MARTINI Force Field. *PLoS Comput. Biol.* **2010**, *6*, e1000810.

(123) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.

(124) Parrinello, M. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182.

(125) Hockney, R.; Eastwood, J. *Computer Simulation Using Particles*; Taylor & Francis, 1988; p 540.

(126) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1-2*, 19–25.

(127) Eisenhaber, F.; Lijnzaad, P.; Argos, P.; Sander, C.; Scharf, M. The Double Cubic Lattice Method: Efficient Approaches to Numerical Integration of Surface Area and

Volume and to Dot Surface Contouring of Molecular Assemblies. *J. Comput. Chem.* **1995**, *16*, 273–284.

(128) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

(129) Anderson, J. A.; Lorenz, C. D.; Travesset, A. General Purpose Molecular Dynamics Simulations Fully Implemented on Graphics Processing Units. *J. Comput. Phys.* **2008**, *227*, 5342–5359.

(130) Glaser, J.; Nguyen, T. D.; Anderson, J. A.; Lui, P.; Spiga, F.; Millan, J. A.; Morse, D. C.; Glotzer, S. C. Strong Scaling of General-Purpose Molecular Dynamics Simulations on GPUs. *Comput. Phys. Commun.* **2015**, *192*, 97–107.

(131) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.

(132) 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.

(133) Raeburn, J.; Cardoso, A. Z.; Adams, D. J. The Importance of the Self-Assembly Process to Control Mechanical Properties of Low Molecular Weight Hydrogels. *Chem. Soc. Rev.* **2013**, *42*, 5143–5156.

(134) Sasselli, I. R.; Halling, P. J.; Ulijn, R. V.; Tuttle, T. Supramolecular Fibers in Gels can be at Thermodynamic Equilibrium: A Simple Packing Model Reveals Preferential Fibril Formation versus Crystallization. *ACS Nano* **2016**, *10*, 2661–2668.

(135) Wang, J.; Liu, K.; Xing, R.; Yan, X. Peptide Self-assembly: Thermodynamics and Kinetics. *Chem. Soc. Rev.* **2016**, *45*, 5589–5604.

(136) Yan, Y.; Huang, J.; Tang, B. Z. Kinetic Trapping–a Strategy for Directing the Self-Assembly of Unique Functional Nanostructures. *Chem. Commun.* **2016**, *52*, 11870–11884.

(137) Mann, S. Self-Assembly and Transformation of Hybrid Nano-objects and Nanostructures under Equilibrium and Non-Equilibrium Conditions. *Nat. Mater.* **2009**, *8*, 781.

(138) Modler, A.; Gast, K.; Lutsch, G.; Damaschun, G. Assembly of Amyloid Protofibrils via Critical Oligomers–A Novel Pathway of Amyloid Formation. *J. Mol. Biol.* **2003**, *325*, 135–148.

(139) Rubinstein, M.; Colby, R. *Polymer Physics*; Oxford University Press: Oxford, 2003; p 113.

(140) Sedghi, M.; Goual, L.; Welch, W.; Kubelka, J. Effect of Asphaltene Structure on Association and Aggregation Using Molecular Dynamics. *J. Phys. Chem. B* **2013**, *117*, 5765–5776.

(141) Wattis, J. A. An Introduction to Mathematical Models of Coagulation–Fragmentation Processes: A Discrete Deterministic Mean-field Approach. *Phys. D* **2006**, *222*, 1–20.

(142) Leyvraz, F. Scaling Theory and Exactly Solved Models in the Kinetics of Irreversible Aggregation. *Phys. Rep.* **2003**, *383*, 95–212.

(143) Treat, R. P. An Exact Solution of the Discrete Smoluchowski Equation and its Correspondence to the Solution of the Continuous Equation. *J. Phys. A* **1990**, *23*, 3003–3016.

(144) Leyvraz, F.; Tschudi, H. R. Singularities in the Kinetics of Coagulation Processes. *J. Phys. A* **1981**, *14*, 3389–3405.

(145) Olivier, B. J.; Sorensen, C. M. Variable Aggregation Rates in Colloidal Gold: Kernel Homogeneity Dependence on Aggregant Concentration. *Phys. Rev. A* **1990**, *41*, 2093–2100.

(146) Olivier, B. J.; Sorensen, C. M.; Taylor, T. W. Scaling Dynamics of Aerosol Coagulation. *Phys. Rev. A* **1992**, *45*, 5614–5623.

(147) Smoluchowski, M. Drei Vortrage uber Diffusion, Brownsche Bewegung und Koagulation von Kolloidteilchen. *Z. Phys.* **1916**, *17*, 557–585.

(148) Smoluchowski, M. Versuch einer Mathematischen Theorie der Koagulationskinetik Kolloider lösungen. *Pisma Mariana Smoluchowskiego* **1927**, *2*, 595–639.

(149) Grassberger, P.; Procaccia, I. Measuring the Strangeness of Strange Attractors. *Phys. D* **1983**, *9*, 189–208.

(150) Wang, J.; Ferguson, A. L. Mesoscale Simulation of Asphaltene Aggregation. *J. Phys. Chem. B* **2016**, *120*, 8016–8035.

(151) Theiler, J. Efficient Algorithm for Estimating the Correlation Dimension from a Set of Discrete Points. *Phys. Rev. A* **1987**, *36*, 4456–4462.

(152) Grzybowski, B. A.; Fitzner, K.; Paczesny, J.; Granick, S. From Dynamic Self-Assembly to Networked Chemical Systems. *Chem. Soc. Rev.* **2017**, *46*, 5647–5678.

(153) Onsager, L. The Effects of Shape on the Interaction of Colloidal Particles. *Ann. N. Y. Acad. Sci.* **1949**, *51*, 627–659.

(154) Chipot, C.; Pohorille, A. *Free Energy Calculations*; Springer, 2007.

(155) Parzen, E., Tanabe, K., Kitagawa, G., Eds. *Selected Papers of Hirotugu Akaike*; Springer Series in Statistics; Springer New York: New York, NY, 1998.

(156) Salvador, S.; Chan, P. Determining the Number of Clusters/Segments in Hierarchical Clustering/Segmentation Algorithms. 16th IEEE International Conference on Tools with Artificial Intelligence. 2004; pp 576–584.

(157) Arora, J. *Introduction to Optimum Design*, 4th ed.; Academic Press: Oxford, UK, 2011.

(158) Reches, M.; Gazit, E. Casting Metal Nanowires within Discrete Self-Assembled Peptide Nanotubes. *Science* **2003**, *300*, 625–627.

(159) Ling, J.; Hutchinson, M.; Antono, E.; Paradiso, S.; Meredig, B. High-Dimensional Materials and Process Optimization Using Data-Driven Experimental Design with Well-Calibrated Uncertainty Estimates. *Integrating Materials and Manufacturing Innovation* **2017**, *6*, 207–217.

(160) Dehghannasiri, R.; Xue, D.; Balachandran, P. V.; Yousefi, M. R.; Dalton, L. A.; Lookman, T.; Dougherty, E. R. Optimal Experimental Design for Materials Discovery. *Comput. Mater. Sci.* **2017**, *129*, 311–322.

(161) Xue, D.; Xue, D.; Yuan, R.; Zhou, Y.; Balachandran, P. V.; Ding, X.; Sun, J.; Lookman, T. An Informatics Approach to Transformation Temperatures of NiTi-based Shape Memory Alloys. *Acta Mater.* **2017**, *125*, 532–541.

(162) Brochu, E.; Cora, V. M.; De Freitas, N. A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning. *arXiv preprint arXiv:1012.2599* **2010**,

(163) Elsaid, L.; Mahamadi, A. Bayesian Optimization for Molecular Communication Channel Capacity. Basic Sciences and Engineering Studies (SGCAC), 2016 Conference of. 2016; pp 183–186.

(164) Griffiths, R.-R. Constrained Bayesian Optimization for Automatic Chemical Design. *arXiv preprint arXiv:1709.05501* **2017**,