Evolution of pi-peptide self-assembly: from understanding to prediction and control

Andrew L. Ferguson[1*] and John D. Tovar[2*]

[1] Pritzker School of Molecular Engineering, University of Chicago, Chicago, Illinois 60637, USA
Email:
andrewferguson@uchicago.edu

[2] Department of Chemistry, Johns Hopkins University, 3400 N. Charles St., Baltimore, MD, 21218 USA.
Email:
tovar@jhu.edu

Abstract

Supramolecular materials derived from the self-assembly of engineered molecules continue to garner tremendous international interest. Recent innovations include the realization of nano- and meso-scale particles (0D), rods and fibrils (1D), sheets (2D) and even extended lattices (3D). Our research groups have focused attention over the past fifteen years on one particular class of supramolecular materials derived from oligopeptides with embedded pi-electron units, where the oligopeptides can be viewed as substituents or side chains to direct the assembly of the central pi-electron cores. Upon assembly, the pi-systems are driven into close co-facial architectures that facilitate a variety of energy migration processes within the nanomaterial volume, including exciton transport, voltage transmission and photoinduced electron transfer. Like many practitioners of supramolecular materials science, many of our initial molecular designs were designed with substantial inspiration from biologically occurring self-assembly coupled with input from chemical intuition and molecular modeling and simulation. In this article, we summarize our current understanding of the pi-peptide self-assembly process as documented through our recent publications in this area. We address fundamental spectroscopic and computational tools used to extract information regarding the internal structures and energetics of the pi-peptide assemblies, and we address the current state of the art in terms of recent applications of data science tools in conjunction with high-throughput computational screening and experimental assays to guide efficient traversal of the pi-peptide molecular design space. The abstract image details our integrated program of chemical synthesis, spectroscopic and functional characterization, multi-scale simulation, and machine learning which has advanced understanding and control of the assembly of synthetic pi-conjugated peptides into supramolecular nanostructures with energy and biomedical applications.

Introduction

Hybrid molecules that capture the optoelectronic functionality of pi-conjugated oligomers with the self-assembly propensities of small oligopeptides have captured the attention of materials researchers at the interface of bio, nano and electronic for well over two decades. This attention could be conceived to have originated from the early compositional introductions of Chmielewski,[1] Bäuerle[2] and Nowick[3] who joined peptide sequences with pi-electron cores. This attention is currently manifest in the many modern day examples showcasing the structures, morphologies and properties of these hybrid bio-electronic materials.[4-19] In general, pi-peptide conjugates can self-assemble into hierarchical nanostructures such as tapes and fibrils in line with many standard oligopeptides that are capable of forming amyloid-like nanomaterials. Importantly, they offer the additional ability to foster intermolecular pi-stacking of the electronic cores which is requisite for energy migration in a broad sense. Thus, pi-peptide structures offer the additional prospects of organic electronics as packaged within a biocompatible vehicle suitable for aqueous deployment.

Our work in this area was first reported in 2008, when Diegelmann *et al*. described the synthesis of a bithiophene-based "amino acid" that could be incorporated into oligopeptide chains via solid-phase peptide synthesis thus leading to a peptide-pi-peptide triblock architecture.[20] In 2010, Vadehra *et al*. described a serendipitous finding whereby biselectrophiles such as diacids or dianhydrides could essentially cross-link resin bound oligopeptides via double acylation or imidation.[21] In 2012, Sanders *et al*. reported the preparation of longer pi-electron components via a piece-wise synthesis approach using Pd-catalysis to join pi-groups within the peptide cores.[22] These synthetic developments allowed us to readily tune the nature of the internal pi-electron core as well as the sequence and directionality of the presented oligopeptide sequences expressed on the core. The impact of these structural variations on the nature of the intermolecular electronic coupling that evolved was assessed via careful interrogation of the resulting electronic properties (via UV-vis, circular dichroism and photoluminescence

spectroscopy) and morphological outcomes (AFM and TEM). We were not alone in pursuing these pi-peptide triblock architectures, as several concurrent studies were published shortly thereafter.[23-25]

Following our initial explorations of sequence and pi-unit structure space, several limitations became clear. First was the lack of incisive computational corroboration for both the electronic coupling outcomes and for the observed nanostructures, both of which require intimate knowledge of the internal intermolecular structure. We were in essence trying to piece together this internal structure by way of simplistic energy minimization protocols. Second was a recognition of the limitations imposed by pure intuitive or Edisonian material design. Our initial peptide sequence selections were based on inspiration from known amyloid-forming peptides, and we expanded into simplistic consensus sequences derived from chemical intuition. These simple sequences worked in a variety of cases, and we made only minor modifications around this local minimum rather than seek out more aggressive shake-ups in structure space.

Almost ten years ago, the authors initiated a collaboration as connected by mutual collaborator Dr. William Wilson, who at the time was with one of us (A.L.F) at the University of Illinois following a prior appointment interacting with one of us (J.D.T.) at Johns Hopkins University. This connection would prove to be quite productive as evidenced by several joint publications and research grants and a continuingly expanding collaboration . Many fruits of this research – both to expand on the sequence and function of the pi-peptide materials and to understand their assembly – have been documented in this journal over the past decade. In this Feature Article, we will highlight our current understanding of how to understand, predict, manipulate and control self-assembly and electronic processes within pi-peptide frameworks.

Pi-peptide primer: structure, assembly, morphology, electronic outcomes

The molecular structure of our peptide-pi-peptide triblock architecture consists of two primary constituents: a central pi-conjugated core structure and two flanking oligopeptide sequences. For our purposes, three to five amino acid residues suffice.[26] We demonstrated the viability of a rich diversity of central pi-cores, spanning hole-transporting oligothiophenes,[22] electron transporting perylene diimides,[21] polymerizable diacetylenes[27] and fluorescent oligophenylene vinylenes. Three- and four- fold peptide functionalization is possible on discotic 2D pi-electron systems.[28] In principle, any suitably functionalized pi-core should be amenable to peptide functionalization provided the core is tolerant of the peptide synthesis chemistry. The specific peptide sequences are similarly quite vast, where individual residues can be selected in order to impart a particular solubility, assembly propensity or bioactivity.

These molecules can be driven to self-assemble similar to standard oligopeptides. In our pi-peptides, we typically rely on acidic amino acid residues such as aspartic or glutamic acid to confer solubility of the pi-peptide molecules under basic pH due to the deprotonation and solvation of the carboxylic acids presented on the pi-peptide (left vial in Figure 1a). Under acidic conditions, the charge screening of the carboxylates provokes extended self-assembly into a variety of fibrillar nanostructures (Figure 1b). In the most basic representation, we can envision the assembly to occur via extended hydrogen bonding networks between the peptides in geometries that enable pi-stacking between the embedded pi-cores, as illustrated simplistically in Figure 1c. The outcome of this process is generally the formation of 1-D nanostructures that maintain structural dimensions of the starting pi-peptide, although bundled structures under ca 10—15 nm are commonly observed via AFM and TEM. Macroscopically, at sufficient concentration, this supramolecular polymerization leads to the formation of self-supporting hydrogels due to the entanglement of the resulting 1-D nanostructures (right vial in Figure 1a).

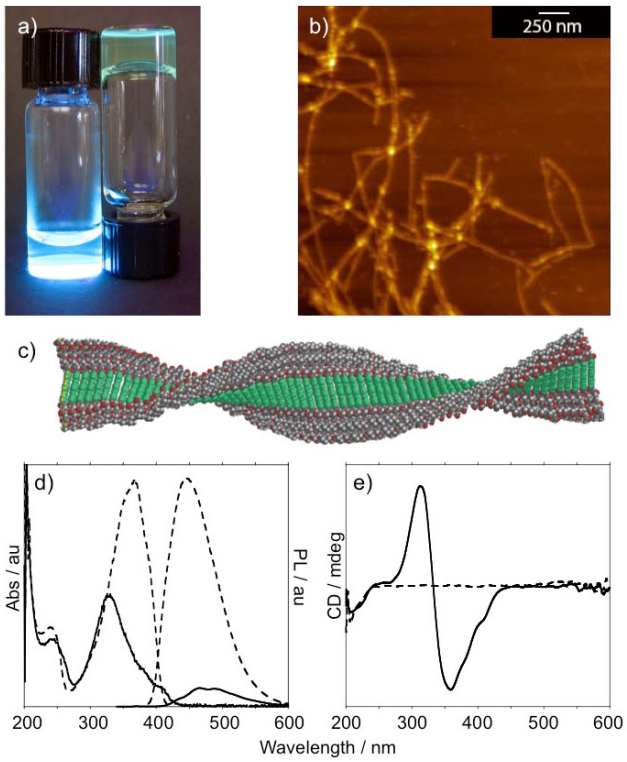


Figure 1. Macro to micro depictions of pi-peptide self-assembly. Vials of pi-peptides in water illuminated with a hand-held UV lamp (a) under basic conditions (at left) and under acidic conditions (at right) showing the self-supporting hydrogel formation in the latter. AFM visualization (b) of the 1-D nanostructures that form in the acidic medium. An energy-minimized model (c) of the resulting assembled fibril-like 1-D structures. UV-vis and PL (d) and CD (e) spectra highlight the optoelectronic changes that occur upon the transition from basic pH (dotted lines) where the pi-peptides are expected to be essentially molecularly dissolved to acidic pH (solid lines) where they are assembled and driven into exciton coupled organization. Reprinted with permission from Ref. 29. Copyright 2015 John Wiley and Sons.

The self-assembly process is also accompanied by general spectroscopic changes (UV-vis, photoluminescence, circular dichroism, Figure 1d-e) that reveal the electronic nature of the internal organization within the resulting nanostructures. Although peptide structure can be grossly interrogated via peptide n-pi* interactions among the amide bonds, pi-peptides have an

added element of pi-core pi-pi* transitions which depend on the nature of the internal chromophore. Under conditions that promote molecular dissolution (typically, basic pH to deprotonate any carboxylates), the spectral properties of a particular pi-peptide are simply those of the embedded pi-core. There is negligible circular dichroism response associated with the lower energy pi-pi* transitions (at 300 nm and higher) due to the lack of molecular chirality of the pi-core. However, upon self-assembly, important spectral details can be extracted which are associated with the phenomenon of exciton coupling as the transition dipoles of the internally embedded pi-cores come into closer proximity. This important aspect is generally manifest in a blue-shifted absorption profile, a quenched photoluminescence, and a pronounced CD response associated with the intermolecular interactions of pi-cores in the chiral aggregates formed during peptide-driven assembly.

The specific processes and outcomes of this assembly vary depending on the nature of the internal pi-conjugated core, and as such they will not be discussed in depth here. Interested readers can gain more insight about these foundational processes in our prior review articles specific to these pi-peptide materials.[26, 29] Suffice it to say that these important spectral observables underlie the important functional outcomes of this assembly in terms guiding a variety of energy transfer applications within these nanomaterials,[30] spanning exciton energy transfer,[31] carrier mobility[32] and electron-transfer behavior.[33-34] Rather than describe case-by-case examples, in this article we will describe our evolution in thinking about how to understand and control the pi-peptide assembly process. This will take us on a journey through experimental and computational design, many of the results of which have been reported in ***Langmuir*** over the past decade.

Sequence variation of electronic coupling

After our initial synthetic developments led us to general routes to include different pi-electron core structures, we next sought to learn how to vary the electronic couplings within the nanomaterials by way of peptide modifications. The ease of solid-phase peptide synthesis allowed us to achieve subtle tuning of the "side-chains" of the pi-electron units in quick order. We anticipated that this could be used as a proxy for the side-chain engineering of alkyl chains that also have a dramatic impact on pi-core solid-state organization.[35] Starting from a DFAG tetrapeptide sequence identified in our early work as being useful to promote pi-peptide assembly, we systematically varied the amino acids in the three positions closest to the pi-core thus leading to three libraries of pi-peptides: DFAX, DFXG and DXAG, where X denotes the point of residue variation (that is, directly next to, one residue removed and two residues removed, respectively, from the central pi core, Figure 2).[36] In basic pH, these libraries all generally showed the same spectral outcome, as expected for the same chromophore in essentially molecular isolation. However, upon acidification, dramatic differences were observed. For example, in the DFAX library, we noted that larger bulky hydrocarbon residues (isoleucine, phenylalanine) led to very strong electronic coupling, while smaller residues (glycine) led to disordered excimer-like states, both as determined by the stark differences in photoluminescence features between these two extreme states (Figure 2a). These assignments are inspired by seminal experimental and theoretical studies on crystalline oligomeric pi-electron aggregates.[37-39] Although there was a nice trend for the DFAX series, it was actually counter to what we anticipated intuitively: we expected that the larger residues would lead to less coupling due to steric packing concerns. To complicate matters, DFXG and DXAG libraries did not present any clear trends, a demonstration that there is no "one size fits all" explanation of structure-function space in these seemingly simple pi-peptides.

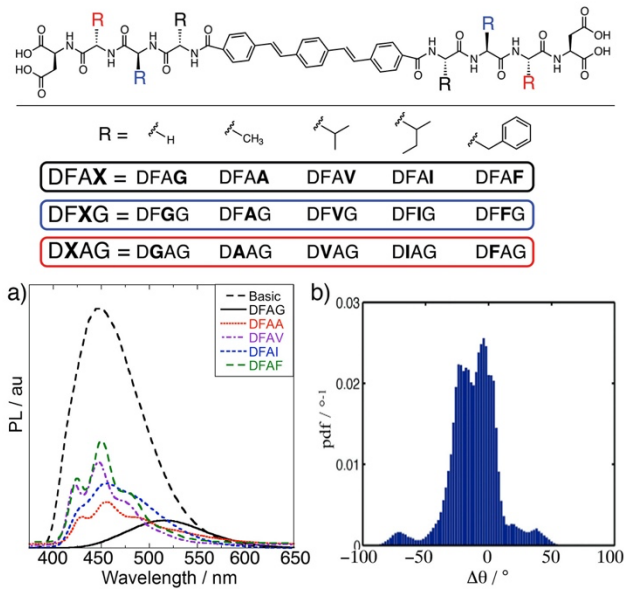


Figure 2. Permutations of an OPV-based pi-peptide bearing four amino acid residues on each peptide flanking wing led to a series of "site-directed mutants" (top) with specific alteration of residue size and hydrophobicity at different positions with respect to the central OPV core. UV-vis spectroscopy (a) revealed different extents of coupling strengths as manifested in strong vibronic features (DFAV, purple trace), broad excimeric features (DFAG, black trace) and combinations thereof (DFAA, red trace). Computational studies show a distribution in the probability density function (pdf) over twist angles Δθ between adjacent peptides (b) following relaxation from an organized assembly with an initial offset of 10° between peptides in a 1-D stack. Adapted with permission from Ref. 36. Copyright 2014 American Chemical Society.

It was at this point that the Baltimore team made the tag to bring the (then) Urbana-Champaign team to the mat. To gain insight into the observed experimental trends, we performed our first molecular modeling studies of pi-peptide families to better understand the molecular behaviors underpinning the structural and functional properties of the self-assembled nanoaggregates. Our initial computational forays employed all-atom classical molecular dynamics simulations to predict the spacing and twist angles between nearest-neighbor pi-peptides in 1D

nanoaggregates as a function of peptide sequence (e.g. Figure 2b).[36] Focusing initially on the DFAX, DFXG and DXAG families, our modeling results demonstrated that larger amino acid residues tended to result in closer intermolecular packing of the pi-peptides within the supramolecular aggregates. This finding was consistent with, and provided a molecular level rationalization for, the experimentally measured photophysical trends: bulky residues placed close to the pi-core produce enhanced dispersion interactions that lead to closer intermolecular spacings and enhance exitonic couplings between pi-core transition dipoles.

The success of these initial computational studies in providing molecular-level rationalization of experimental trends and in guiding new modalities of design thinking in our rational engineering efforts led to a rapid expansion in the computational effort and seeded a robust experimental/computational collaboration that came to define our subsequent investigations of these materials. To date, our computational program has incorporated all-atom, coarse-grained, and ultra-coarse-grained classical molecular simulations, first principles density functional theory calculations, and machine learning-enabled design strategies employing deep representational learning, high-throughput virtual screening, Bayesian optimization, and hybrid experimental/computational active learning strategies. We touch on many of these aspects throughout the remainder of this review.

Based on our synthetic strategies, another structural variable at our disposal was the sequence directionality of the peptide backbone with respect to the central pi-electron core. Our initial designs had one uniform N-to-C directionality, and thus, one N terminus and one C terminus for a particular pi-peptide. Later synthetic innovations simplified the preparative chemistry but now directed the peptides to emanate from the pi-core in the same sense, that is, providing two C termini to the pi-peptide molecule. In principle, the peptides with the single N and C termini, dubbed “non-symmetric” peptides, would need to form a tighter anti-parallel beta-sheet type of structure in order to achieve intimate intermolecular pi-electron coupling, while “symmetric”

peptides with two C termini would need to form a parallel beta-sheet to achieve the same type of coupling. We posited that these differences would be manifest in the extent of exciton coupling, so we prepared two peptides based on the same DFAG residue presentation relative to a central oligophenylene vinylene core, but that varied in the sense of the C-to-N directionality (Figure 3 top).[40] From a first glance, clear differences were indeed noted: the symmetric case with the less organized hydrogen-bonding network provided signatures associated with a disordered excimer state (Figure 3a) while the non-symmetric case showed evidence for more tightly coupled chromophores (Figure 3b). However, direct comparisons are complicated by the fact that the electronic nature of the central pi-core in these two molecules slightly differs: the symmetric pi-peptide bears two carboxamide substituents while the non-symmetric version has one carboxamide and one N-acyl substituent.

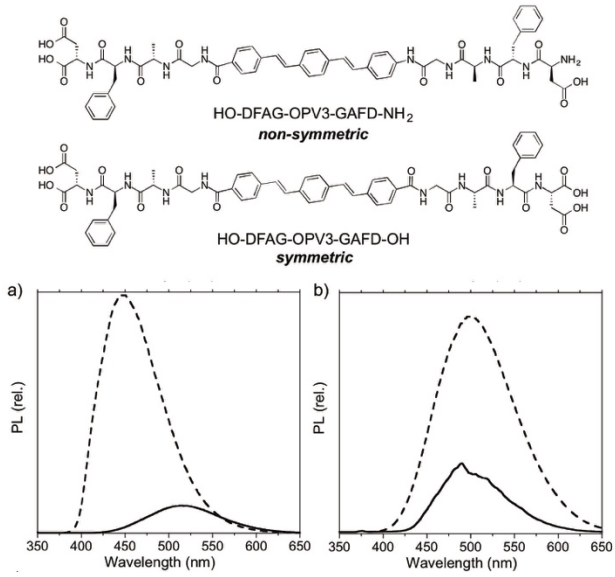


Figure 3. Depictions of non-symmetric and symmetric OPV pi-peptides (top) which yielded different excited state outcomes upon assembly as revealed by PL spectroscopy (solid traces in a and b, respectively). The dotted lines reflect the dissolved pi-peptide in basic pH. Adapted with permission from Ref. 40. Copyright 2014 American Chemical Society.

All-atom, coarse-grained, and patchy particle models of pi-peptide assembly

Motivated by a desire to gain understanding of the molecular structures adopted by the various dimers described in the previous section, we conducted all-atom molecular dynamics simulations employing enhanced sampling techniques to calculate the structure and free energy of the dimer as a function of the pi-core chemistry and orientation of the peptide wings.[40] Our calculations revealed that the thermodynamic favorability of dimer formation could be strongly tuned by the directionality of the peptide wings, and that for a smaller pi-core the non-symmetric peptides produced more strongly bound dimers, whereas for a larger pi-core this trend inverted. A decomposition of the calculated free energy of dimerization into its constituent energetic and entropic components also proved valuable in advancing understanding of the molecular driving forces mediating assembly. Somewhat unexpectedly at the time, we found dispersion interactions to be a primary mediator of this interaction, accounting in some cases for as much as 79% of the dimerization interaction energy. This finding stood contrary to prior assumption and expectation that hydrogen bonds were the stabilizing interaction of paramount importance and this result guided our design thinking towards engineering dispersion interactions to control assembly.

Building upon this work, we conducted more detailed characterization of the thermodynamics and kinetics of multimerization of DFAG-OPV3-GAFD peptides containing an oligo-phenylenevinylene (OPV) core.[41] With regards to thermodynamics, all-atom calculations in explicit solvent revealed a favorable dimerization free energy of $\Delta F \approx -15\ k_BT$ and an even more favorable and approximately constant free energy of association for additional monomers of $\Delta F \approx -25\ k_BT$. Decomposition of the free energy into its constituent contributions revealed that dispersion and Coulombic interactions distributed across both the pi-cores and peptide wings play a significant role in mediating the enthalpy of dimer formation.[40,41] Interestingly, even in the high-pH state in which the four carboxyl groups within the C-terminal Asp residues are deprotonated and the pi-peptides carry a (-4) net charge, we predicted the formation of small oligomers containing several

molecules to be thermodynamically spontaneous driven by favorable dispersion interactions. This led to the computational prediction that pi-peptides should exist as multimers under high-pH conditions, as opposed to monomers stabilized by Coulombic repulsion as previously thought. This prediction was subsequently experimentally validated using dynamic light scattering and fluorescence correlation spectroscopy.[42] This result updated our understanding of the acid-triggered assembly of elongated nanoaggregates to proceed not by simple monomeric addition, but by the agglomeration of pre-existing oligomers.

Turning to kinetics, we used our explicit solvent calculations to perform bottom-up parameterization of an implicit solvent model that enabled us to reach longer time and length scales and directly simulate the self-assembly of hundreds of molecules over tens of ns on commodity compute clusters (Figure 4a).[41] By fitting a continuous time Markov chain (CTMC) to the data, we constructed a predictive model for the short-time ($\lesssim$100 ns) assembly kinetics. The predictions of the CTMC model were in good agreement with the simulation data and predicted a hierarchical assembly mechanism wherein light aggregates reorganize and assemble into ordered β-sheet-like stacks that subsequently agglomerate into larger disordered aggregates with internal structural relaxation time scales exceeding several tens of ns. Interestingly, this assembly mechanism is reminiscent of that previously reported for amyloid fibrils that share substantial chemical similarity with pi-peptides.[43-48]

Our development of a computationally inexpensive implicit solvent model motivated us to construct additional models at increasing degrees of coarse-grained resolution to probe even larger length and time scales in order to directly observe the later stages of assembly. We employed the popular Martini force field[49-50] to develop a bead-level coarse-grained model for the DFAG-OPV3-GAFD peptide and used Boltzmann inversion[51] to recalibrate the bonded interactions against all-atom, explicit solvent simulation data.[52] The resulting coarse-grained model was validated against all-atom distribution functions and free energy profiles, and then

used to conduct simulations of hundreds of pi-peptides over hundreds of ns. In pleasing consistency with the predictions of our implicit-solvent all-atom calculations, we indeed find assembly at high-pH from an initial monomeric dispersion produces a distribution of oligomers containing up to eight molecules over the course of ~300 ns. Assembly is self-limiting due to the buildup of negative charge within the growing oligomer that thermodynamically disfavors further growth. At low-pH, we also observed good consistency with our prediction of a hierarchical assembly mechanism and the irreversible production of well-aligned dimers, trimers, and tetramers that form the building blocks for larger and more disordered ellipsoidal and branched aggregates (Figure 4b). These larger structures ultimately grow to incorporate all molecules in the system and undergo internal structural ripening on time scales exceeding hundreds of ns. The large size of our simulation cell enabled us to perform structural characterization of the self-assembled branched nanoaggregates to show that they are fractal networks with fractal dimensionality of ~1.5. Fitting of a Smoluchowski coagulation model[53] to the irreversible aggregation kinetics revealed assembly to be self-similar with cluster size in the limit of long times and large cluster sizes, weakly dependent on the initial conditions, and possess a consistent assembly mechanism that does not change with concentration. This provided deep insight into the fundamental molecular mechanisms underpinning pi-peptide assembly and the existence of an emergent simplicity in the dynamics of assembly that are well-described by a simple analytical model with a single adjustable parameter. In follow-up work, we further interrogated this coarse-grained model to provide molecular understanding of the observation that slow acidification can produce more ordered self-assembled aggregates[54] and show that shear flows used to guide assembly in experimental protocols only influence the orientation and aggregation of clusters at length scales of hundreds of monomers over time scales of hundreds of nanoseconds.[55]

To further push the envelope of accessible length and time scales further, we developed an ultra-coarse-grained patchy particle model capable of comfortably simulating ten thousand pi-

peptides over time scales of hundreds of microseconds and length scales of hundreds of nanometers.[56] Contrary to the bottom-up parameterization scheme adopted in the development of the implicit solvent and Martini bead-level models, in this case we adopted a top-down strategy. We designed a rigid particle formed by the union of several spheres designed to approximately mimic the excluded volume and direction interactions of a generic DXXX-Π-XXXD pi-peptide, and specified appropriate parameter ranges based on all-atom and Martini models of molecules containing experimentally accessible pi-cores and the full range of amino acids in each flanking peptide sequence. Our simulations revealed the formation of a porous fractal network at length scales ≳30 nm (tens of peptide lengths) with fractal dimensionality of ~2 regardless of the particular choice of model parameters (Figure 4c). The branches within this network correspond to linear supramolecular aggregates containing tens of stacked peptides, and the network nodes to the intersections of these pseudolinear aggregates. The morphology is reminiscent of experimentally observed “matted hair” structures formed by pi-peptides under quiescent conditions.[57] Indeed, hydrodynamic forces are frequently employed in experimental assembly protocols to break up this network and align the strands into long pseudolinear nanoaggregates comprising ca. 1-2 thousand stacked molecules per micron.[27, 58] At shorter length scales between 1 nm and 30 nm corresponding to the length scale of single fibrillar branches, we observe a fractal dimensionality that depends on the particular model parameters but – in good agreement with our Martini-based bead-level model – is centered on ~1.5. We also performed Pareto optimization of the model to discover parameter regimes leading to the formation of pseudolinear aggregates with maximally ordered pi-core stacking – a prerequisite for high mobility intermolecular charge transport – and identify specific sequences within the DXXX-OPV3-XXXD family consistent with these parameter values. In addition to providing new understanding of the large-scale structure of the self-assembled aggregates, the model was used to prospectively identify promising new peptide sequences not previously explored in experiment. This application of our molecular

models for predictive sequence identification was a turning point in our work that led us to integrate our molecular models into high throughput virtual screening pipelines, data-driven modeling paradigms, and active learning loops to guide rational traversal and filtration of the vast pi-peptide sequence space to identify the most promising candidates for experimental synthesis and testing. We discuss aspects of these more recent efforts below.

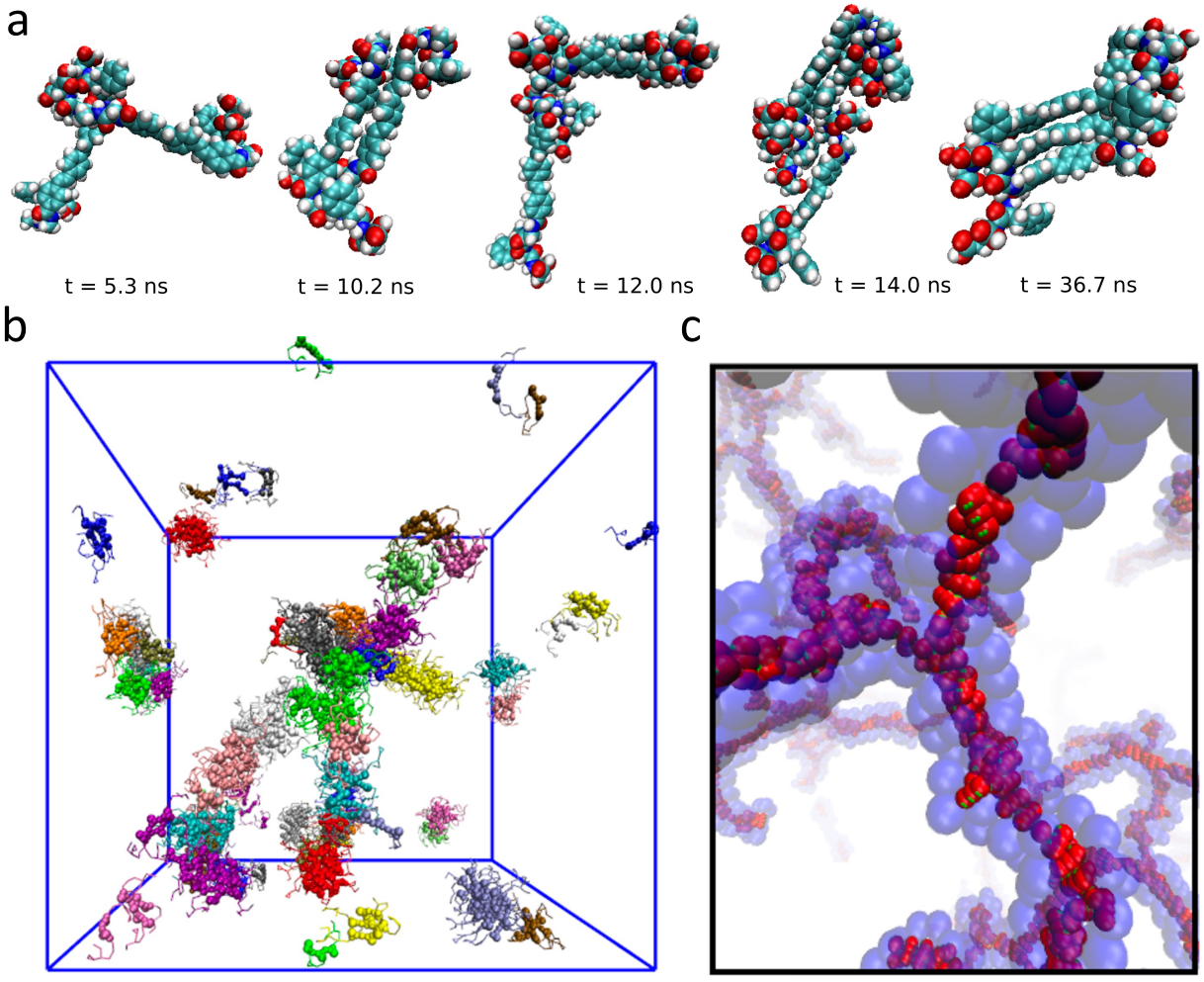


Figure 4. Computational screening, discovery, and design of pi-peptides using molecular modeling at multiple scales. (a) Implicit solvent molecular dynamics simulations enable the direct observation of multimerization events of DFAG-OPV3-GAFD pi-peptides over tens of nanoseconds to reveal the molecular mechanisms of peptide oligomerization and structural

ripening. (Image adapted from Ref. 41: Thurston et al. “Thermodynamics, morphology, and kinetics of early-stage self-assembly of π-conjugated oligopeptides” Mol. Sim. 42 12 955-975 (2016) doi:10.1080/08927022.2015.1125997, reprinted by permission of the publisher Taylor & Francis Ltd.) (b) A bead-level coarse-grained model enables simulation of the assembly of hundreds of DFAG-OPV3-GAFD pi-peptides over hundreds of nanoseconds to characterize the formation and structure of branched aggregates with a fractal dimension of ~1.5. (Adapted with permission from Ref. 52: Mansbach and Ferguson “Coarse-Grained Molecular Simulation of the Hierarchical Self-Assembly of π-Conjugated Optoelectronic Peptides” J. Phys. Chem. B 121 1684–1706 (2017) doi:10.1021/acs.jpcb.6b10165. Copyright 2017 American Chemical Society.) (c) A patchy particle model enables the simulation of ten thousand DXXX-Π-XXXD pi-peptides over hundreds of microseconds enabling observation of the formation of a large-scale porous fractal network reminiscent of the “matted hair” aggregates in experimental assembly in quiescent solvent. (Adapted with permission from Ref. 56: Mansbach and Ferguson “Patchy Particle Model of the Hierarchical Self-Assembly of π-Conjugated Optoelectronic Peptides” J. Phys. Chem. B 122 10219–10236 (2018) doi:10.1021/acs.jpcb.8b05781. Copyright 2018 American Chemical Society.)

Alkyl dilution and peptide field effects

Humanity has divined complex explanations from simple observations for centuries, turning to tea leaves, scattered bones, rune stones and palm lines for insight. Our interpretation of spectral signatures could be similarly classified, as the complex phenomena associated with pi-electron interactions belie the seeming simplicity associated with spectral features. For example, in the aggregation of pi-electron units, there are two major operative processes that lead to the observed spectral outcome: exciton coupling, which for an H-type (cofacial) aggregate

expected within our pi-peptides should lead to a blue-shift in absorption, and planarization, which due to the enhanced intramolecular pi-conjugation should lead to a red-shift. In some of our early examples, the spectral signatures associated with tightly coupled states actually fell within the spectral envelope of the "molecularly dissolved" chromophore, thus leading to some ambiguity in our assessments. To further interrogate these competing influences, we posited that a non-conjugated alkyl group could serve as an electronically silent proxy for the central pi-conjugated core of our pi-peptides,[59] allowing us to probe the electronic changes the pi-peptide undergoes during the self-assembly process but without the associated exciton coupling.

For this study, we prepared two different pi-peptides, one relatively hydrophilic (D: R = CH_2COOH, Figure 5) and one relatively hydrophobic (V: R = $CH(CH_3)_2$, Figure 5). Both pi-peptides were separately co-assembled with the corresponding inert alkyl peptide at different ratios. The expectation was that the pi-peptides would be the minority component of a statistical co-assembly, thus effectively being electronically decoupled from each other. The hydrophilic pi-peptide showed spectral properties consistent with this dilution: co-assembled aggregates presented UV-vis and PL signatures suggesting an isolated OPV unit that coincided with the spectra from the molecularly dissolved peptide at basic pH (blue cartoons in Figure 5). On the other hand, the hydrophobic peptide showed evidence for self-sorting leading to apparent "block-like" supramolecular polymers showing distinct evidence for exciton coupling among the OPV units (red cartoons in Figure 5), even at 5 mol% dilution within the electronically silent matrix. These findings were corroborated by circular dichroism studies revealing that the pi-cores in the hydrophilic co-assemblies were isolated in chiral environments as expected from peptide assembly (black dotted line in Figure 5a), but that the pi-cores in the hydrophobic co-assemblies were actually undergoing exciton coupling due to the self-sorting (black dotted line in Figure 5b). This finding was further challenged in an oligothiophene pi-peptide similarly diluted but with internally di-methylated thiophenes that further contorted the planarity of the oligothiophene at the

molecular level.[60] The torsional constraint impacted the extent of conjugation and coupling within the diluted nanostructures, in a manner that could be tuned via thermal annealing as a way to further influence the optoelectronic properties.

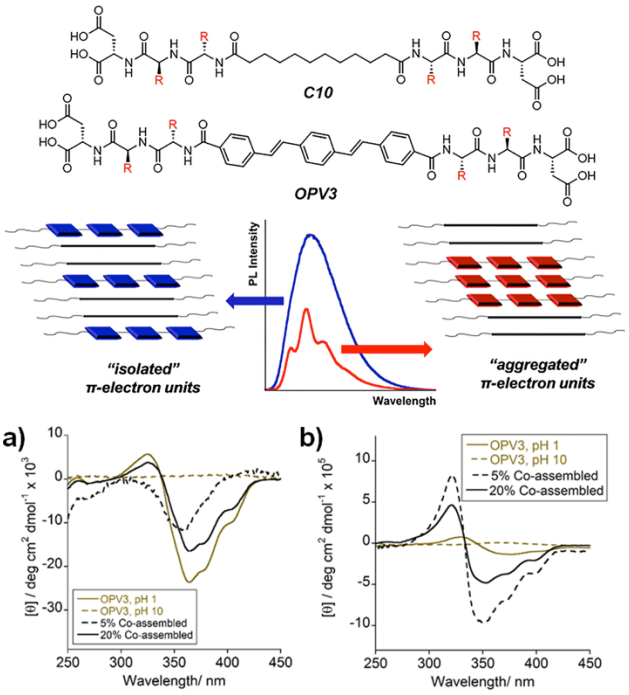


Figure 5. Hydrophilic (DDD) and hydrophobic (DVV) pi-peptides and inert alkyl peptides were used for dilution studies. The DDD pi-peptides as minority components showed statistical dilution (blue cartoons) while the DVV pi-peptides showed evidence for self-sorting into "blocks" (red cartoons) as evidenced through CD spectroscopy (a and b respectively). The wavelength scale for the central PL spectral plots is ca. 375-650 nm. Adapted with permission from Ref. 30 (Copyright 2018 Institute of Physics) and Ref. 59 (Copyright 2017 American Chemical Society).

Variation of internal structure: alkyl spacers

A final example showcasing the impact of new molecular variations comes from our work to install alkyl spacers between the central pi-cores and the flanking peptide sequences.[61] As alluded to above, the nature of the synthesis strategy dictates the constitution of the covalent

linker joining peptide to pi. In our symmetric pi-peptides with two C-termini, these linkers are carboxamide functionalities which are anticipated to serve as electron-withdrawing substituents with respect to the relatively electron-rich pi-cores. Therefore, when considering these pi-peptides as bioorganic semiconductor candidates, their hole-transporting ability might be compromised relative to traditional unsubstituted variants. Although we have measured remarkable mobilities for several types of pi-peptide nanostructures, we thought we could push these values even higher by electronically decoupling the electron-withdrawing carboxamide linker from the central pi-core. A simple variation of our solid-phase "on-resin dimerization" synthetic strategy allowed us to do this in short order, where we could include 1, 2, or 3 carbon spacers between the central pi-core and the peptide wings (Figure 6).[61] We targeted these different lengths cognizant of the weak pi-character invoked within methylene units and of the conformational impacts these carbon chains might pose on overall molecular geometry and subsequent ability to organize into supramolecular polymers.

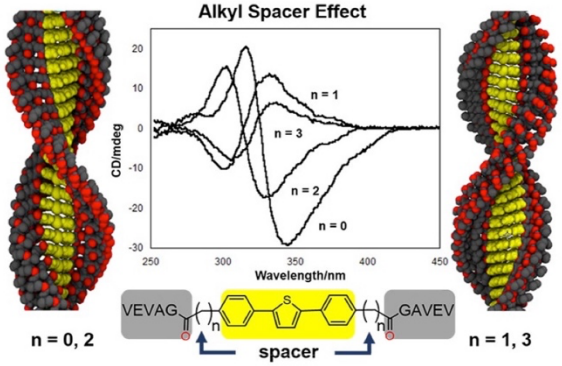


Figure 6. Different helical orientations available to pi-peptides bearing different internal spacer lengths, as revealed by CD spectroscopy. Reprinted with permission from Ref. 61. Copyright 2019 American Chemical Society.

Indeed, the installation of the carbon spacers led to isolated pi-cores not subject to extended conjugation with the carboxamide linkers. Despite different conformational orientations

possible within the different linkers, all pi-peptides still formed 1-D fibrillar nanostructures thus indicating that the basic assembly paradigm was not altered due to alkyl group conformational constraints. Completely unexpected though was the observation of periodic inversion in the observed helicity of the exciton coupling as the carbon length was altered, as observed in the reversal of the Cotton effect in the CD spectra (Figure 6). This alteration followed an "odd-even" trend whereby 0 and 2 carbon spacers yielded M-type helicity (left-handed twists) while 1 and 3 carbon spacers yielded P-type helicity (right-handed twists). This observation held for a variety of central pi-cores which themselves presented different geometric considerations, and they all were appended to the same VEVAG peptide sequence.

To ascertain the underlying molecular determinants of these chiral behaviors, we employed all-atom molecular dynamics simulations of pi-peptide dimers coupled with enhanced sampling techniques to comprehensively sample the range of potential twist angles and interpretable machine learning techniques to isolate the critical determinants of the chiral inversion.[61] Our simulation results were consistent with the experimental observations of P-type preference for odd-length spacers and M-type for even, and our data-driven analysis resolves hydrogen bonding and the steric packing of hydrophobic side chains to be the chief factors dictating supramolecular chirality. This understanding furnished previously unknown design rules and actionable principles by which to rationally engineer chirality into these supramolecular assemblies. For example, our computational analysis indicated that the valine and alanine substituents on a VEVAG pentapeptide that were placed closest to the central pi-core had substantial influence on the self-assembly. To challenge this finding, we systematically swapped the configuration of these centers (using D- and L-amino acids) and probed the resulting chiroptical properties.[62] These changes led to systematic alterations of the expressed macromolecular chirality, whereas changing more distal residues (such as the configuration of

the terminal valine) did not lead to appreciable variation of chirality as observed by circular dichroism.

Predictive modeling

The vast size of chemical space spanned by the X_n-Π-X_n pi-peptides – where X_n denotes a string of n amino acids and Π a pi-conjugated core – means that Edisonian trial-and-improvement experimentation presents an inefficient means to identify the top performing sequences. Chemical intuition and experimental experience are enormously valuable in predicting good candidates for the pi-cores and peptide wings, but computational screening and active learning approaches present the opportunity to perform more principled and efficient exploration of chemical space to identify lead candidates for experimental testing.

We employed a computational screen within the context of our work on the role of the alkyl spacer between the pi-core and peptide wings in governing structural assembly and photophysical properties. We designed the screen to consider VEVAG-Π-GAVEV molecules with six possible pi-cores and alkyl spacers containing 0-3 carbon atoms.[63] We employed a multi-scale computational protocol combining (i) classical all-atom molecular dynamics simulations to compute the emergent structure of the self-assembled nanoaggregates, (ii) first principles density functional theory calculations to predict the stability and mobility of electrons and holes within these aggregates, and (iii) Marcus theory to turn these computed quantities into a prediction of the intermolecular charge mobility. The screen identified top performing molecules with electron mobilities of up to 0.143 $cm^2\ Vs^{-1}$ and hole mobilities of up to 0.224 $cm^2\ Vs^{-1}$. The screen was also useful in furnishing new understanding and design rules. Specifically, we identified particular pi-cores that did and did not favor high charge mobility and determined charge mobility to be anti-correlated with alkyl spacer length suggesting that elimination of the alkyl spacer entirely may enhance charge transport.

Our multi-scale model also enabled us to perform a deeper analysis of the DXX-OT4-XXD pi-peptide family containing a quaterthiophene pi-core to gain insight into the interplay between molecular and electronic structure.[64] Molecular dynamics simulations revealed that smaller amino acid residues promoted linear stacking of neighboring molecules, whereas bulkier ones tended to induce larger relative twist angles. Density functional theory calculations showed the absorption spectrum to be dominated by transitions between carbon and sulfur p orbitals and relatively insensitive to the twist angle. In contrast, the highest occupied molecular orbital (HOMO) was found to be quite sensitive to the relative twist, changing from shared between the two neighboring molecules at small angles (i.e., for linear, in-register stacking of the pi-cores) to becoming strongly localized on one molecule at larger angles and hindering intermolecular charge transport. This study provided actionable principles to engineer charge transport by selecting peptide sequences that produce linear stacking and good sharing of the HOMO between neighboring pi-peptide cores.

We also deployed computational screening to develop a surrogate model linking pi-peptide sequence to the quality of alignment of the pi-cores within the emergent self-assembled nanoaggregates.[65] Focusing on the DXXX-Π-XXXD family with perylene diimide (PDI) and naphthalene diimide (NDI) pi-cores, we constructed a surrogate quantitative structure-property relationship (QSPR) model linking pi-peptide sequence to oligomerization thermodynamics (i.e., dimerization and trimerization free energies) and alignment quality. The model was trained over 26 point mutants of the DFAG-Π-GAFD molecule and constructed in physically motivated and inexpensive physicochemical descriptors. We then used the surrogate model to conduct a high throughput virtual screen (HTVS) of 9826 candidates within the DXXX-Π-XXXD design space and identify top performing candidates with good predicted alignment. In doing so, we identified DAVG-PDI-GVAD as a promising and previously unstudied pi-peptide that we subsequently validated in molecular dynamics simulations. The model also exposed an optimum in the

dimerization and trimerization free energies that maximally promotes the formation of well-aligned nanoaggregates, presenting an actionable design rule for pi-peptide sequence engineering.

Based on this computational insight, we prepared this DAVG PDI peptide along with a DAIA PDI peptide, another sequence identified as fostering high co-facial intermolecular orientation.[66] As predicted, the DAVG sequence yielded spectroscopic properties indicative of a strongly cofacial organization, while the DAIA sequence showed evidence for less structured stacking that is common for PDI aggregates (Figure 7). We hypothesized that, in line with our prior work on the DXXX peptides discussed above, the alanine and glycine residues most proximal to the central pi-core were playing a major role in dictating the assembly outcomes. We therefore designed a new set of peptides VEVXG and VEVXA whereby the X residue was a small screen of "hydrophobic" residues (Gly, Ala, Val, Phe). VEVXG peptides showed behavior comparable to DAVG while VEVXA peptides were comparable to DAIA. This highlights a general structure feature of the proximal amino acid residue to impact PDI aggregate photophysics in these pi-peptide nanomaterials and reveals the importance of computational screening to identify new leads from which intuitive extensions can be made.

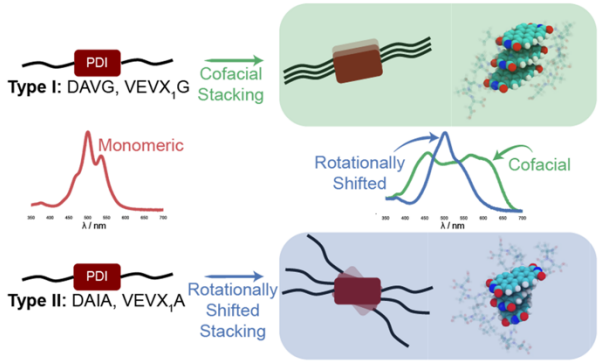


Figure 7. Two distinct regimes of electronic coupling that were revealed with the help of computational screening. DAVG pi-peptides (and others bearing proximal glycine residues) showed greater preferences for cofacial packing (top) while DAIA pi-peptides (and others bearing

proximal alanine residues) showed more shifted coupling (bottom). Reprinted with permission from Ref. 66. Copyright 2021 American Chemical Society.

A deficiency of one-shot virtual screens is that there is no provision for feedback and iteration wherein the results of the screen are used to prospectively identify putatively superior new sequences to incorporate into the screen. Active learning strategies present a powerful means to build a data-driven model over the space of molecules screened to date and use this model to guide rational traversal of the design space by identifying promising new candidates. These candidates are then screened, the model updated, and additional candidates proposed within a virtuous cycle: the model accuracy improves with each round and is therefore better able to prospectively identify top performing candidates. The process is terminated when a sufficient number of candidates meeting the design criteria are identified or no further improvements are observed over a number of iterative rounds. We integrated an active learning protocol with our Martini-based bead-level model to screen the 8000 members of the DXXX-OPV3-XXXD family of pi-peptides containing an oligophenylenevinylene pi-core.[67] Our surrogate model linked the peptide wing sequence to the degree of assembly of pseudolinear nanoaggregates observed in molecular simulations. We used a variational autoencoder[68] to provide a learned low-dimensional featurization of the molecules and Gaussian process regression[69] to construct a quantitative model linking sequence to nanoaggregate structure. We then closed the active learning loop by interrogating the regression model using Bayesian optimization[70] to select the next batch of sequences for molecular simulation. We seeded the active learning screen with simulations of 90 randomly selected sequences from the design space, and then conducted 25 rounds of active learning sampling up to four molecules per round. After sampling a total of 186 candidates comprising just 2.3% of the total design space, we ceased to see any additional improvements in the model or its predictions and terminated the screen. We used the terminal model to make a

priority ranking of all 8000 DXXX-OPV3-XXXD molecules that was both consistent with existing physicochemical understanding of which amino acid residues lead to good assembly behavior and identified novel sequences predicted to exhibit superior assembly behaviors to those previously reported.

Very recently, we employed both computation and experiments within an active learning paradigm (Figure 8).[71] In doing so, we have incorporated experimentation directly within the active search rather than just as a post hoc test of the computational predictions. This paradigm is extremely powerful in enabling the surrogate models to learn and receive feedback from high-fidelity experimental assays while simultaneously benefiting from the more comprehensive sampling of design space accessible to high throughput but more approximate (i.e., lower fidelity) computational models. We applied this hybrid computational/experimental active learning screen to the design space of 694,982 X_n-4T-X_n pi-peptides containing a quaterthiophene pi-core and flanked by symmetric peptide wings of up to five amino acids.[71] We employed graph-based message passing neural networks within a regularized autoencoding architecture to perform deep representational learning of a low-dimensional featurization of the pi-peptide design space,[72] and then constructed surrogate models linking sequence to computational and experimental performance measures using multi-objective[70, 73] and multi-fidelity[74] Gaussian process regression. We performed 38 rounds of computation and three rounds of experiments within the active learning screen to sample 1181 molecules by computation and 28 by experiment and efficiently identify several experimentally validated sequences that were completely novel and with assembly performance on par with known top performing candidates. Although the precise ratio can fluctuate depending on the computational models employed and the experimental analyses conducted, it is typical that computational screening loops can operate approximately 1-2 orders of magnitude faster than experimental loops. In this particular case, computational screening of a single oligopeptide required approximately 48 hours of wall clock time – although many could be screened simultaneously using parallel

compute resources – whereas the experimental preparation and assay of a single molecule required several days and had to be typically performed in serial. Interpretation of our model predictions also exposed a number of comprehensible design rules for high performing sequences. For example, the model identified a statistically significant trend that (D/E)X_n(A/G) motifs tend to promote good pi-core alignment whereas the incorporation of presence of large hydrophobes such as Leu, Ile, Phe, and Tyr disfavor good assembly. These trends are consistent with our prior understanding that placing hydrophilic and polar groups far from the pi-core and small non-polar groups adjacent to the pi-core maintains the hydrophobic environment in the proximity of the core necessary for good stacking without introducing too much disruptive steric bulk. Similarly, incorporating bulky hydrophobes anywhere within the sequence is consistent with our prior observations that these residues can participate in hydrophobic interactions with the pi-cores to compete with and disrupt core-core stacking[61]. Interestingly, the model also suggested that proline residues – almost completely unexplored in our pi-peptide studies – may promote the formation of elongated nanoaggregates with good pi-core stacking. We currently lack a mechanistic understanding for this prediction, but note that recent experimental work by Müllen, Bäuerle, Wennemers, and co-workers demonstrated pi-conjugated peptide assembly employing proline peptide wings attached to quaterthiophene cores.[15]

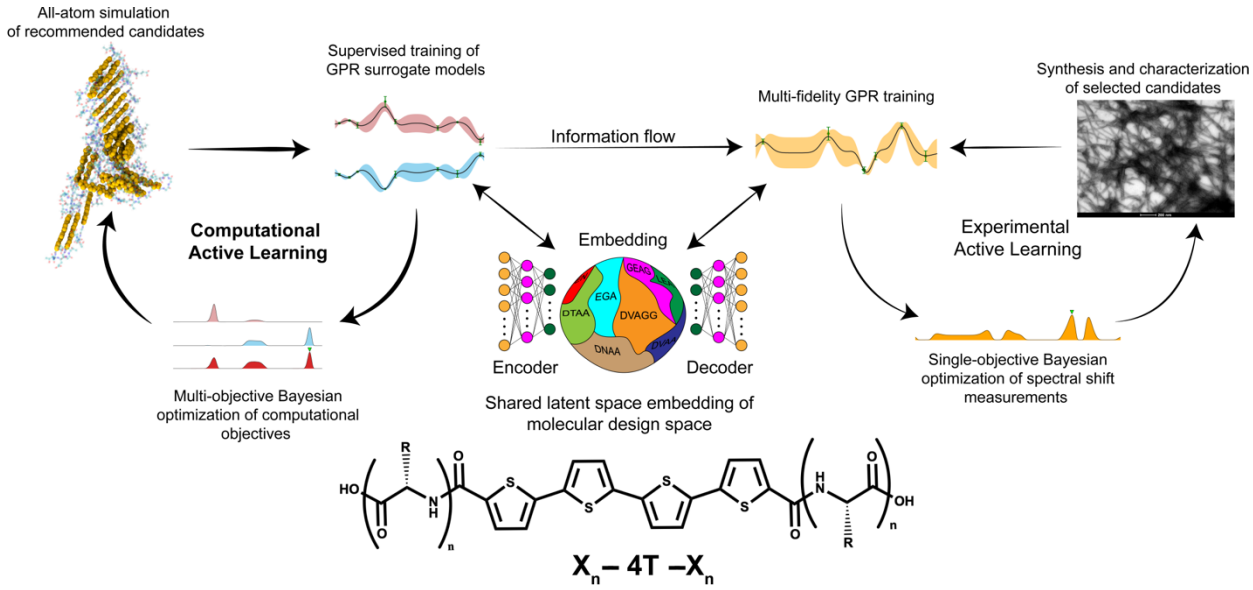


Figure 8. Illustration of the hybrid computational/experimental multi-fidelity active learning loop for the data-driven design of X_n-4T-X_n pi-peptides containing a quaterthiophene pi-core and flanked by symmetric peptide wings of up to five amino acids. High-throughput computation and low-throughput experiment are integrated within a multi-fidelity surrogate model to predict the assembly performance as a function of the amino acid sequence of the peptide wings. The hybrid model optimally combines accurate but sparse experimental measurements with voluminous but approximate computational predictions to guide efficient exploration of the pi-peptide design space. We interleaved 38 rounds of computation and three rounds of experiment in which we computationally simulated 1181 candidates and experimentally assayed 28 candidates to discover several novel sequences with assembly performance commensurate with known top performing pi-peptides. (Reproduced from Ref. 71 with permission from the Royal Society of Chemistry.)

Conclusions and outlook

Synthetic oligopeptides with embedded pi-electron cores present versatile building blocks for the self- and directed-assembly of supramolecular nanoaggregates with emergent electronic

and optical functionality. Specific applications of these nanomaterials include energy harvesting and transport within biological environments, interfacing between biotic and abiotic systems, and as novel biocompatible semiconducting substrates. Through an integrated decade-long collaborative program of experimental synthesis and characterization, computational modeling, and machine learning, we uncovered fundamental principles governing the sequence-structure-function relationships underpinning these molecular materials and developed integrated experimental and computational processes and pipelines to efficiently traverse molecular design space and identify candidate molecules with engineered supramolecular structure and function. Particular accomplishments of this program of work include: the establishment of design rules for amino acid placement within the peptide wings to promote in-register stacking and intermolecular electronic coupling of pi-cores within high-aspect ratio nanoaggregates; the development of multi-scale computational models to gain molecular understanding spanning from electronic structure to large-scale fractal networks observable by microscopy; understanding and control of alkyl spacers between the pi-core and peptide wings to modulate supramolecular chirality; and active learning pipelines to ingest computational predictions and experimental assays and parameterize data-driven surrogate models to prospectively identify molecular candidates and guide exploration of pi-peptide design space. These achievements have enabled us to wield unprecedented levels of control over the supramolecular structure and function of these pi-peptides, but a number of challenging scientific and technical questions remain and exciting frontiers un(der)explored. For example, how can we interface these molecules with other biological constituents like proteins, lipids, or nucleic acids to generate hybrid nanostructures? Can we use these molecules to perform high-efficiency exciton splitting in a device? Can we use pi-peptide nanostructures to guide and direct biomineralization? Can we develop automated “self-driving laboratories” to discover new molecules orders of magnitudes faster than the decade-long research program reviewed herein?

We perceive a bright future for pi-peptides as powerful molecular building blocks with which to engineer biocompatible nanomaterials with electronic and optical functionality. Based on our collective synthetic experience, spectral measurement and computational assessment, we developed many sophisticated energy management schemes, consisting of field effect behavior, excitonic energy migration and photoinduced electron transfer. These schemes in some cases required co-assembly approaches whereby pi-peptides with different pi-core energy offsets could be manipulated into a variety of complex nanomaterials. As a first step towards biomedical applications, and in line with other contemporary efforts to explore therapeutic applications of pi-peptide constructs in general,[75] we initiated in vitro cell studies showcasing how pi-peptide hydrogel mechanics can influence neurite extension among human neural stem cells.[76] We also explored the first in vivo deployment of these types of triblock peptide-pi-peptide molecules in the context of hydrogel formation in the brain following stroke,[77] where we could include low-energy dyes to facilitate fluorescence imaging without complications from biological autofluorescence. Our future efforts will entail understanding how the electric field creation within these nanomaterials can further influence in vitro cell physiology or other in vivo outcomes. We envision these materials as potentially serving as a next generation substrate for energy harvesting, transport, and semi-conducting applications in biological environments, and we believe the coming years will capitalize on foundational molecular understanding to drive both new basic science and translational applications.

Acknowledgements. This material reviews work supported by the U.S. Department of Energy Office of Basic Energy Sciences under Grant Nos. DE-SC0004857 and DE-SC0011847 and the National Science Foundation under Grant Nos. DMR-1729011, DMR-1841807, DMR-1728947, DMR-1407493 and DGE-1746045. Computational components of the work were completed in

part with resources provided by the National Energy Research Scientific Computing Center (NERSC), a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231, 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 and which is a joint effort of the University of Illinois at Urbana–Champaign and the National Center for Supercomputing Applications, the University of Illinois Computational Science and Engineering Program, and the University of Chicago Research Computing Center. We gratefully acknowledge computing time on the University of Chicago high-performance GPU-based cyberinfrastructure supported by the National Science Foundation under Grant No. DMR-1828629. Parts of this work were performed while A.L.F. was visiting the Institute for Pure and Applied Mathematics (IPAM), which is supported by the National Science Foundation under Grant No. DMS-1440415. We thank Dr. Ben Blaiszik for his assistance in hosting simulation data on the Materials Data Facility. Finally, we acknowledge the important intellectual contributions of all co-authors whose work is described here.

Conflicts of Interest. A. L. F. is a co-founder and consultant of Evozyne, Inc. and a co-author of US Patent Application 16/887,710, US Provisional Patent Applications 62/853,919, 62/900,420, and 63/314,898 and International Patent Applications PCT/US2020/035206 and PCT/ US2020/050466.

References

(1) Ulysse, L.; Chmielewski, J., The Synthesis of a Light-Switchable Amino-Acid for Inclusion into Conformationally Mobile Peptides. *Bioorg. Med. Chem. Lett.* **1994,** *4*, 2145-2146.
(2) 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.

(3) Gothard, C. M.; Rao, N. A.; Nowick, J. S., Nanometer-sized amino acids for the synthesis of nanometer-scale water-soluble molecular rods of precise length. *J. Am. Chem. Soc.* **2007,** *129*, 7272-7273.
(4) Heinz-Kunert, S. L.; Pandya, A.; Dang, V. T.; Tran, P. N.; Ghosh, S.; McElheny, D.; Santarsiero, B. D.; Ren, Z.; Nguyen, A. I., Assembly of pi-Stacking Helical Peptides into a Porous and Multivariable Proteomimetic Framework. *J. Am. Chem. Soc.* **2022,** *144*, 7001-7009.
(5) Wei, D.; Yu, Y. Z.; Ge, L. L.; Wang, Z. F.; Chen, C.; Guo, R., Chiral Supramolecular Polymers Assembled by Amphiphilic Oligopeptide-Perylene Diimides and High Electrochemical Sensing. *Langmuir* **2021,** *37*, 9232-9243.
(6) Tian, T.; Wei, D.; Ge, L. L.; Wang, Z. F.; Chen, C.; Guo, R., Hierarchical self-assemblies of carnosine asymmetrically functioned perylene diimide with high optoelectronic response. *J. Coll. Interface Sci.* **2021,** *601*, 746-757.
(7) Peressotti, S.; Koehl, G. E.; Goding, J. A.; Green, R. A., Self-Assembling Hydrogel Structures for Neural Tissue Repair. *ACS Biomater. Sci. Eng.* **2021,** *7*, 4136-4163.
(8) Kaur, H.; Roy, S., Enzyme-Induced Supramolecular Order in Pyrene Dipeptide Hydrogels for the Development of an Efficient Energy-Transfer Template. *Biomacromolecules* **2021,** *22*, 2393-2407.
(9) Fortunato, A.; Sanzone, A.; Mattiello, S.; Beverina, L.; Mba, M., The pH- and salt-controlled self-assembly of 1 benzothieno 3,2-b 1 -benzothiophene-peptide conjugates in supramolecular hydrogels. *New J. Chem.* **2021,** *45*, 13389-13398.
(10) Boddula, R.; Singh, S. P., Peptide-based novel small molecules and polymers: unexplored optoelectronic materials. *J. Mater. Chem. C* **2021,** *9*, 12462-12488.
(11) Rani, A.; Kavianinia, I.; Hume, P.; De Leon-Rodriguez, L. M.; Kihara, S.; Williams, D. E.; McGillivray, D. J.; Plank, N. O. V.; Gerrard, J.; Hodgkiss, J. M.; Brimble, M. A., Directed self-assembly of peptide-diketopyrrolopyrrole conjugates - a platform for bio-organic thin film preparation. *Soft Matter* **2020,** *16*, 6563-6571.
(12) Njenga, S. M.; Wang, X.; Jiang, W.; Wan, X. B., Nanostructure Control of a Regioregular Poly(3-alkylthiophene) Using an Oligopeptide Side Chain. *Macromolecules* **2020,** *53*, 6087-6098.
(13) Hazra, S.; Shit, A.; Ghosh, R.; Basu, K.; Banerjee, A.; Nandi, A. K., Modulation of the optoelectronic properties of a donor-acceptor conjugate between a cationic polythiophene and a peptide appended perylene bisimide ampiphile. *J. Mater. Chem. C* **2020,** *8*, 3748-3757.
(14) Hafner, R. J.; Gorl, D.; Sienkiewicz, A.; Balog, S.; Frauenrath, H., Long-Lived Photocharges in Supramolecular Polymers of Low-Band-Gap Chromophores. *Chem. Eur. J.* **2020,** *26*, 9506-9517.
(15) Ochs, N. A. K.; Lewandowska, U.; Zajaczkowski, W.; Corra, S.; Reger, S.; Herdlitschka, A.; Schmid, S.; Pisula, W.; Mullen, K.; Bauerle, P.; Wennemers, H., Oligoprolines guide the self-assembly of quaterthiophenes. *Chem. Sci.* **2019,** *10*, 5391-5396.
(16) Guo, Z. X.; Wang, Y. J.; Zhang, X.; Gong, R. Y.; Mu, Y. B.; Wan, X. B., Solvent-Induced Supramolecular Assembly of a Peptide-Tetrathiophene-Peptide Conjugate. *Front. Chem.* **2019,** *7*.
(17) Egan, J. G.; Brodie, G.; McDowall, D.; Smith, A. J.; Edwards-Gayle, C. J. C.; Draper, E. R., Impact of subtle change in branched amino acid on the assembly and properties of perylene bisimides hydrogels. *Mater. Adv.* **2021,** *2*, 5248-5253.

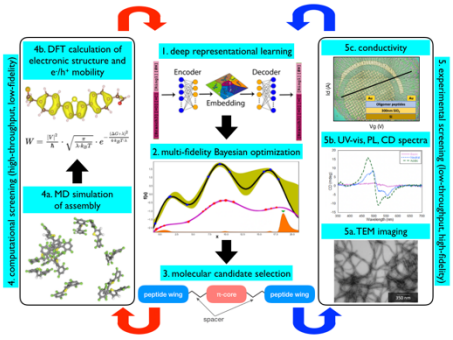
(18) Cross, E. R.; Sproules, S.; Schweins, R.; Draper, E. R.; Adams, D. J., Controlled Tuning of the Properties in Optoelectronic Self-Sorted Gels. *J. Am. Chem. Soc.* **2018,** *140*, 8667-8670.
(19) Ashkenasy, N.; Horne, W. S.; Ghadiri, M. R., Design of self-assembling peptide nanotubes with delocalized electronic states. *Small* **2006,** *2*, 99-102.
(20) 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.
(21) Vadehra, G. S.; Wall, B. D.; Diegelmann, S. R.; Tovar, J. D., On resin dimerization incorporates a diverse array of pi-conjugated functionality within aqueous self-assembling peptide backbones. *Chem. Commun.* **2010,** *46*, 3947-3949.
(22) Sanders, A. M.; Dawidczyk, T. J.; Katz, H. E.; Tovar, J. D., Peptide-based supramolecular semiconductor nanomaterials via facile Pd-catalyzed solid-phase 'dimerizations'. *ACS Macro Lett.* **2012,** *1*, 1326-1329.
(23) Schillinger, E. K.; Mena-Osteritz, E.; Hentschel, J.; Börner, H. G.; Bäuerle, P., Oligothiophene Versus b-Sheet Peptide: Synthesis and Self-Assembly of an Organic Semiconductor-Peptide Hybrid. *Adv. Mater.* **2009,** *21*, 1562-1567.
(24) Stone, D. A.; Hsu, L.; Stupp, S. I., Self-assembling quinquethiophene-oligopeptide hydrogelators. *Soft Matter* **2009,** *5*, 1990-1993.
(25) 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.
(26) Tovar, J. D., Supramolecular Construction of Optoelectronic Biomaterials. *Acc. Chem. Res.* **2013,** *46*, 1527-1537.
(27) Diegelmann, S. R.; Hartman, N.; Markovic, N.; Tovar, J. D., Synthesis and alignment of discrete polydiacetylene-peptide nanostructures. *J. Am. Chem. Soc.* **2012,** *134*, 2028-2031.
(28) Sanders, A. M.; Kale, T. S.; Katz, H. E.; Tovar, J. D., Solid-phase synthesis of self-assembling multivalent pi-conjugated peptides. *ACS Omega* **2017,** *2*, 409-419.
(29) Tovar, J. D., Peptide Nanostructures with π-Ways: Photophysical Consequences of Peptide/π-Electron Molecular Self-Assembly. *Isr. J. Chem.* **2015,** *55*, 622-627.
(30) Tovar, J. D., Photon management in supramolecular peptide nanomaterials. *Bioinspir. Biomim.* **2018,** *13*, 015004.
(31) Ardoña, H. A. M.; Tovar, J. D., Energy transfer within responsive pi-conjugated coassembled peptide-based nanostructures in aqueous environments. *Chem. Sci.* **2015,** *6*, 1474-1484.
(32) Besar, K.; Ardona, H. A. M.; Tovar, J. D.; Katz, H. E., Demonstration of Hole Transport and Voltage Equilibration in Self-Assembled pi-Conjugated Peptide Nanostructures Using Field-Effect Transistor Architectures. *ACS Nano* **2015,** *9*, 12401-12409.
(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.; 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.

(35) Jeffries-EL, M.; McCullough, R. D., Regioregular Polythiophenes. In *Handbook of Conducting Polymers*, 3 ed.; Skotheim, T. A.; Reynolds, J. R., Eds. CRC Press: New York, 2007; Vol. 1, pp 9.1-9.45.
(36) Wall, B. D.; Zacca, A. E.; Sanders, A. M.; Wilson, W. L.; Ferguson, A. L.; Tovar, J. D., Supramolecular Polymorphism: Tunable Electronic Interactions within pi-Conjugated Peptide Nanostructures Dictated by Primary Amino Acid Sequence. *Langmuir* **2014,** *30*, 5946-5956.
(37) Sherwood, G. A.; Cheng, R.; Smith, T. M.; Werner, J. H.; Shreve, A. P.; Peteanu, L. A.; Wildeman, J., Aggregation Effects on the Emission Spectra and Dynamics of Model Oligomers of MEH-PPV. *J. Phys. Chem. C* **2009,** *113*, 18851-18862.
(38) Spano, F. C., The Spectral Signatures of Frenkel Polarons in H- and J-Aggregates. *Acc. Chem. Res.* **2010,** *43*, 429-439.
(39) Gierschner, J.; Ehni, M.; Egelhaaf, H. J.; Milian Medina, B.; Beljonne, D.; Benmansour, H.; Bazan, G. C., Solid-state optical properties of linear polyconjugated molecules: pi-stack contra herringbone. *J. Chem. Phys.* **2005,** *123*, 144914.
(40) 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 pi-conjugated oligopeptide nanostructures. *Langmuir* **2014,** *30*, 11375-11385.
(41) Thurston, B. A.; Tovar, J. D.; Ferguson, A. L., Thermodynamics, morphology, and kinetics of early-stage self-assembly of pi-conjugated oligopeptides. *Mol. Sim.* **2016,** *42*, 955-975.
(42) Valyerde, 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,** *34*, 7346-7354.
(43) Gillam, J. E.; MacPhee, C. E., Modelling amyloid fibril formation kinetics: mechanisms of nucleation and growth. *J. Phys. Condensed Matter* **2013,** *25*, 373101.
(44) Wagoner, V. A.; Cheon, M.; Chang, I.; Hall, C. K., Computer simulation study of amyloid fibril formation by palindromic sequences in prion peptides. *Proteins Struct. Funct. Bioinformatics* **2011,** *79*, 2132-2145.
(45) Schmit, J. D.; Ghosh, K.; Dill, K., What Drives Amyloid Molecules To Assemble into Oligomers and Fibrils? *Biophys. J.* **2011,** *100*, 450-458.
(46) Li, Y.; Roberts, C. J., Lumry-Eyring Nucleated-Polymerization Model of Protein Aggregation Kinetics. 2. Competing Growth via Condensation and Chain Polymerization. *J. Phys. Chem. B* **2009,** *113*, 7020-7032.
(47) Andrews, J. M.; Roberts, C. J., A Lumry-Eyring nucleated polymerization model of protein aggregation kinetics: 1. Aggregation with pre-equilibrated unfolding. *J. Phys. Chem. B* **2007,** *111*, 7897-7913.
(48) Yong, W.; Lomakin, A.; Kirkitadze, M. D.; Teplow, D. B.; Chen, S. H.; Benedek, G. B., Structure determination of micelle-like intermediates in amyloid beta-protein fibril assembly by using small angle neutron scattering. *Proc. Nat. Acad. Sci. USA* **2002,** *99*, 150-154.
(49) 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. *Journal of Chemical Theory and Computation* **2008,** *4*, 819-834.
(50) Marrink, S. J.; Tieleman, D. P., Perspective on the Martini model. *Chem. Soc. Rev.* **2013,** *42*, 6801-6822.

(51) Noid, W. G., Perspective: Coarse-grained models for biomolecular systems. *J. Chem. Phys.* **2013,** *139*.
(52) 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,** *121*, 1684-1706.
(53) Wattis, J. A. D., An introduction to mathematical models of coagulation-fragmentation processes: A discrete deterministic mean-field approach. *Physica D Nonlinear Phenomena* **2006,** *222*, 1-20.
(54) Adams, D. J.; Butler, M. F.; Frith, W. J.; Kirkland, M.; Mullen, L.; Sanderson, P., A new method for maintaining homogeneity during liquid-hydrogel transitions using low molecular weight hydrogelators. *Soft Matter* **2009,** *5*, 1856-1862.
(55) Mansbach, R. A.; Ferguson, A. L., Control of the hierarchical assembly of pi-conjugated optoelectronic peptides by pH and flow. *Organic & Biomolecular Chemistry* **2017,** *15*, 5484-5502.
(56) Mansbach, R. A.; Ferguson, A. L., Patchy Particle Model of the Hierarchical Self-Assembly of pi-Conjugated Optoelectronic Peptides. *Journal of Physical Chemistry B* **2018,** *122*, 10219-10236.
(57) Li, B.; Li, S. S.; Zhou, Y. C.; Ardoña, H. A. M.; Valverde, L. R.; Wilson, W. L.; Tovar, J. D.; Schroeder, C. M., Nonequilibrium Self-Assembly of pi-Conjugated Oligopeptides in Solution. *ACS Appl. Mater. Interfaces* **2017,** *9*, 3977-3984.
(58) 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 the shear flow assembly of peptide hydrogels. *Adv. Mater.* **2011,** *23*, 5009-5014.
(59) Ardoña, H. A. M.; Kale, T. S.; Ertel, A.; Tovar, J. D., Nonresonant and Local Field Effects in Peptidic Nanostructures Bearing Oligo(p-phenylenevinylene) Units. *Langmuir* **2017,** *33*, 7435-7445.
(60) Kale, T. S.; Ardona, H. A. M.; Ertel, A.; Tovar, J. D., Torsional Impacts on Quaterthiophene Segments Confined within Peptidic Nanostructures. *Langmuir* **2019,** *35*, 2270-2282.
(61) Panda, S. S.; Shmilovich, K.; Ferguson, A. L.; Tovar, J. D., Controlling Supramolecular Chirality in Peptide-pi-Peptide Networks by Variation of the Alkyl Spacer Length. *Langmuir* **2019,** *35*, 14060-14073.
(62) Panda, S. S.; Shmilovich, K.; Ferguson, A. L.; Tovar, J. D., Computationally Guided Tuning of Amino Acid Configuration Influences the Chiroptical Properties of Supramolecular Peptide-pi-Peptide Nanostructures. *Langmuir* **2020,** *36*, 6782-6792.
(63) Shmilovich, K.; Yao, Y. F.; Tovar, J. D.; Katz, H. E.; Schleife, A.; Ferguson, A. L., Computational discovery of high charge mobility self-assembling pi-conjugated peptides. *Mol. Sys. Des. Eng.* **2022,** *7*, 447-459.
(64) Thurston, B. A.; Shapera, E. P.; Tovar, J. D.; Schleife, A.; Ferguson, A. L., Revealing the Sequence-Structure-Electronic Property Relation of Self-Assembling pi-Conjugated Oligopeptides by Molecular and Quantum Mechanical Modeling. *Langmuir* **2019,** *35*, 15221-15231.
(65) Thurston, B. A.; Ferguson, A. L., Machine learning and molecular design of self-assembling pi-conjugated oligopeptides. *Mol. Sim.* **2018,** *44*, 930-945.

(66) Panda, S. S.; Shmilovich, K.; Herringer, N. S. M.; Marin, N.; Ferguson, A. L.; Tovar, J. D., Computationally Guided Tuning of Peptide-Conjugated Perylene Diimide Self-Assembly. *Langmuir* **2021,** *37*, 8594-8606.
(67) Shmilovich, K.; Mansbach, R. A.; Sidky, H.; Dunne, O. E.; Panda, S. S.; Tovar, J. D.; Ferguson, A. L., Discovery of Self-Assembling pi-Conjugated Peptides by Active Learning-Directed Coarse-Grained Molecular Simulation. *J. Phys. Chem. B* **2020,** *124*, 3873-3891.
(68) Kingma, D. P.; Welling, M., Auto-Encoding Variational Bayes. *arXiv* **2013**, arXiv:1312.6114.
(69) Rasmussen, C. E.; Williams, C. K. I., *Gaussian Processes for Machine Learning*. The MIT Press: Cambridge, MA, 2005.
(70) Brochu, E.; Cora, V. M.; De Frietas, N., A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint* **2010**, arXiv:1012.2599.
(71) Shmilovich, K.; Panda, S. S.; Stouffer, A.; Tovar, J. D.; Ferguson, A. L., Hybrid computational-experimental data-driven design of self-assembling pi-conjugated peptides. *Digital Discovery* **2022,** *1*, 448-462.
(72) Ghosh, P.; Sajjadi, M. S. M.; Vergari, A.; Black, M.; Schölkopf, B., From Variational to Deterministic Autoencoders. **2019**, arXiv:1903.12436.
(73) Paria, B.; Kandasamy, K.; Poczos, B. In *A Flexible Framework for Multi-Objective Bayesian Optimization using Random Scalarizations*, 35th Uncertainty in Artificial Intelligence (UAI) Conference, Tel Aviv, ISRAEL, Jul 22-25; Tel Aviv, ISRAEL, 2019; pp 766-776.
(74) Perdikaris, P.; Raissi, M.; Damianou, A.; Lawrence, N. D.; Karniadakis, G. E., Nonlinear information fusion algorithms for data-efficient multi-fidelity modelling. *Proceedings of the Royal Society a-Mathematical Physical and Engineering Sciences* **2017,** *473*, **20160751**.
(75) Zou, Q. L.; Abbas, M.; Zhao, L. Y.; Li, S. K.; Shen, G. Z.; Yan, X. H., Biological Photothermal Nanodots Based on Self-Assembly of Peptide Porphyrin Conjugates for Antitumor Therapy. *J. Am. Chem. Soc.* **2017,** *139*, 1921-1927.
(76) Liyanage, W. G. M.; Ardoña, H. A. M.; Mao, H. Q.; Tovar, J. D., Cross-linking approaches to tune the mechanical properties of peptide-pi-electron based hydrogels. *Bioconj. Chem.* **2017,** *28*, 751-759.
(77) Dibble, J. P.; Deboer, S. R.; Mersha, M.; Robinson, T. J.; Felling, R. J.; Zeiler, S. R.; Tovar, J. D., In Vivo Formation and Tracking of pi-Peptide Nanostructures. *ACS Appl. Mater. Interfaces* **2022,** *in press (10.1021/acsami.2c04598).*

TOC Graphic



An integrated program of chemical synthesis, spectroscopic and functional characterization, multi-scale simulation, and machine learning has advanced understanding and control of the assembly of synthetic pi-conjugated peptides into supramolecular nanostructures with energy and biomedical applications.

Biographies (with photos)

Andrew L. Ferguson is an Associate Professor and Vice Dean for Equity, Diversity, and Inclusion at the Pritzker School of Molecular Engineering at the University of Chicago. He received an M.Eng. in Chemical Engineering from Imperial College London in 2005, a Ph.D. in Chemical and Biological Engineering from Princeton University in 2010, and from 2010 to 2012 he was a Postdoctoral Fellow of the Ragon Institute of MGH, MIT, and Harvard in the Department of Chemical Engineering at MIT. He commenced his faculty career in Materials Science and Engineering at the University of Illinois at Urbana-Champaign in 2012 and joined the Pritzker School of Molecular Engineering in 2018. His research uses theory, simulation, and machine learning to design self-assembling biomaterials, understand macromolecular folding, develop antiviral therapies, and establish enhanced sampling techniques for molecular simulation.

John. D. Tovar is a Professor of Chemistry and of Materials Science and Engineering at Johns Hopkins University. Following undergraduate studies in Chemistry (UCLA), graduate studies in Organic Chemistry (MIT) and postdoctoral work in Materials Science (Northwestern), he rose through the ranks at Johns Hopkins, where his research interests lie in the exploration of pi-electron materials with unusual manifestations of aromaticity and conjugation, and the exploration

of supramolecular pi-electron interactions as mediated through peptide substituents. Photo credit: Will Kirk, JHU.